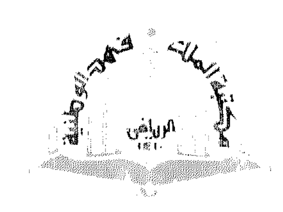
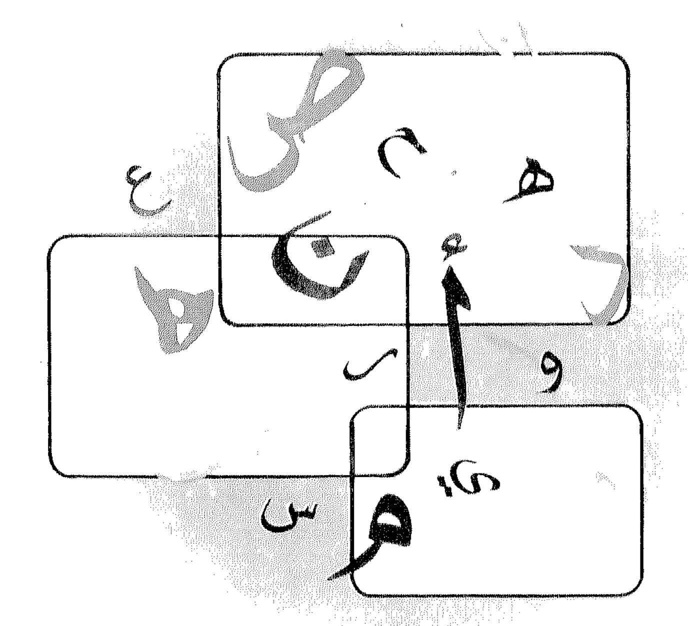
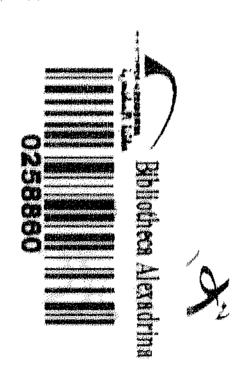




مطبوعات
مكتبة الملك فهد الوطنية
السلسلة الثانية
(٢١)



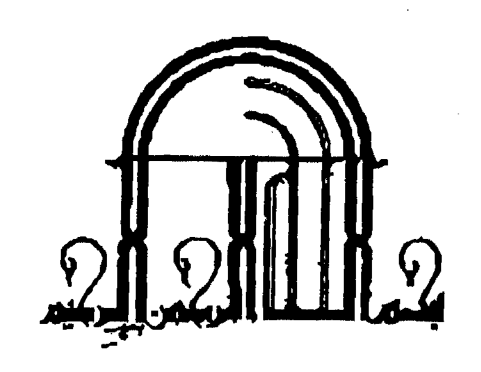
في اللغة العربية



د. [illegible]مان الصوينع

الرياض
١٤١٥هـ / ١٩٩٤م


Converted by Tiff Combine - (no stamps are applied by registered version)


بسم الله الرحمن الرحيم

مطبوعات
مكتبة الملك فهد الوطنية
السلسلة الثانية (٢١)

تهتم هذه السلسلة بنشر الدراسات والبحوث
في إطار علم المكتبات والمعلومات بشكل عام

استرجاع المعلومات
في اللغة العربية

علي السليمان الصوينع

الرياض
١٤١٥هـ / ١٩٩٤م

فهرسة مكتبة الملك فهد الوطنية

الصوينع ، علي السليمان
استرجاع المعلومات في اللغة العربية .
١٧٦ص ؛ ٢٤ سم ؛ (السلسلة الثانية ؛ ٢١)
ردمك ٠-٣٧-٠٠-٩٩٦٠
ردمد ١٣١٩-٢٩٨١
١ - اختزان واسترجاع المعلومات - اللغة العربية
٢- المكتبات - معالجة بيانات ٣- البرمجة (حاسبات إلكترونية) أ - العنوان ب - السلسلة

ديوي ٠٢٧,٧ ١٥/٠٥٧٧

رقم الإيداع :١٥/٠٥٧٧
ردمك : ٠-٣٧-٠٠-٩٩٦٠
ردمد : ١٣١٩-٢٩٨١

المحتويات

الموضوع	الصفحة
تصدير	٩
الفصل الأول (نظم الاسترجاع في المكتبات)	١١
طبيعة نظم الاسترجاع في المكتبات	١٣
متطلبات تصميم نظام الاسترجاع	١٧
أدبيات الموضوع	٢٣
الفصل الثاني (لغويات المعلومات)	٢٩
النص	٣١
المفهوم	٣٤
أشكال المفاهيم	٣٥
المفاهيم المفردة	٣٦
المفاهيم المركبة	٣٩
الأسماء	٤٥
المصطلحات الأجنبية	٥١
اللغة الطبيعية والاصطناعية	٦٠
خصائص الاسترجاع باللغة الطبيعية	٦٣
مشكلات الاسترجاع باللغة الطبيعية	٦٨

التصريف والاشتقاق ٦٨
المترادفات ٧٠
الألفاظ المشتركة ٧٠
التركيب ٧٢
صيغة الفعل ٧٤
الإملاء ورسم الكلمات ٧٥
الرمز والمجاز ٧٦
التحقيق والاستدعاء ٧٧
الفصل الثالث (كلمات التوقف) ٧٩
خصائص كلمات التوقف ٨٤
أنواع كلمات التوقف ٨٨
الحروف والأدوات النحوية ٨٩
المفردات غير الدالة ٩٤
تكوين كلمات التوقف ١١٠
الفصل الرابع ("أل" التعريف والاسترجاع) ١١٧
مشكلة (أل) الأصلية ١٢٢
الحروف المتصلة ١٢٥
الفصل الخامس (الترتيب والمواصفات) ١٣١
الترتيب الألفبائي ١٣٣
المحارف العربية ١٣٦

التحليل الصرفي ١٣٩
التحليل الصرفي ونظم المعلومات ١٤٢
بتر الكلمات ١٤٣
الفصل السادس (التباديل والاسترجاع) ١٤٩
كشافات التباديل ١٥١
عرض المداخل ١٥٤
الترتيب والاسترجاع ١٥٨
مداخل العناوين ١٥٩
مداخل الأسماء ١٦٣
المستخلصات ١٦٦
المراجع ١٧١



تصدير

سبق أن نشر هذا الكتاب في طبعته الأولى عام ١٩٨٨م تحت عنوان (كشافات التباديل واسترجاع المعلومات في اللغة العربية) . وقد مضى على صدور هذا الكتاب سنوات كثيرة طرأت خلالها تطورات غنية ومتنوعة في مجالات التقنية واسترجاع المعلومات خصوصًا مع بداية انتشار الأقراص البصرية لتحميل النصوص والمصورات الوثائقية على نطاق واسع .

أما على مستوى نظم المعلومات فقد طرح العديد من البرامج والنظم الجيدة في معالجة واسترجاع المعلومات للعديد من التطبيقات باللغة العربية ، إلا أن أغلب الحلول المطروحة قد لا تفي بالاحتياجات الكاملة والمميزة لطبيعة البحث والاسترجاع في مجالات المكتبات ومراكز المعلومات التى يحاول هذا الكتاب دراستها من وجهة نظر الموثقين والمكتبيين بالدرجة الاولى .

ومنذ صدور هذا الكتاب في طبعته الأولى توالت المؤتمرات والبحوث والمقالات حول العديد من القضايا النظرية والتطبيقية التي تتناول جوانب مختلفة لاستخدام اللغة العربية في تقنية الحاسوب وبالذات ما له صلة مباشرة باسترجاع المعلومات والتكشيف الآلي ، أو لغويات المعلومات في حقل المكتبات . وقد رأيت استتمامًا للفائدة الإشارة إلى البحوث الحديثة ذات العلاقة في سياق الحديث عن لغويات المعلومات .

وتأتي الرغبة في إعادة طباعة الكتاب مرة أخرى بعد نفاده من جهة ، بالاضافة إلى ازدياد الاهتمام بدراسة مشكلات نظم استرجاع المعلومات باللغة العربية من جهة أخرى ، هذا علاوة على تنامي التطبيقات والبرامج المطروحة باللغة العربية .

والكتاب في الأساس لا يتناول قضايا متغيرة مع تطور النظم والتقنية وإنما يعالج مسائل أساسية أو تطبيقية في مجالات اللغة العربية والتوثيق والاسترجاع ، ولذا فإن جل الزيادات أو التعديلات التي طرأت على الطبعة الثانية موجهة نحو التوسع في الجوانب العلمية التي تهم العاملين فى المكتبات ومراكز المعلومات عند استخدام الحاسوب لمعالجة واسترجاع المعلومات باللغة العربية ، كما يتطرق الكتاب إلى بعض القضايا والقواعد والأساليب التي تهم الباحثين والدارسين في مجالات لغويات المعلومات ونظم الاسترجاع وتطبيقاتها باللغة العربية بشكل عام .

وحيث إن موضوعات الكتات تجاوزت نطاق موضوع كشافات التباديل سواء في الأصل أم مع الإضافات الكثيرة في الطبعة الثانية ، فقد رأيت تغيير عنوان الكتاب بحيث يصبح (استرجاع المعلومات في اللغة العربية) .

ويمكن استعراض الموضوعات العامة والمخصصة التي يتناولها الكتاب في المداخل التالية :

* استرجاع المعلومات - الحاسوب - كشافات التباديل - اللغة العربية - لغويات المعلومات - المكتبات ومراكز المعلومات .
* أل التعريف - بتر الكلمات - التحليل الصرفي - الترتيب الألفبائي - الحروف العربية - العناوين العربية - كلمات التوقف - الكلمات الدالة - اللغة الطبيعية - اللغة الاصطناعية - المحارف العربية - المستخلصات .

والله أسأل أن يهدينا سبيل الرشاد وأن يعم بنفعه العباد.

الرياض : ١٤١٥هـ

١٩٩٤م

علي سليمان الصوينع

الفصل الأول

نظم الاسترجاع في المكتبات

- طبيعة نظم الاسترجاع في المكتبات
- متطلبات تصميم نظام الاسترجاع
- أدبيات الموضوع



الفصل الاول
نظم الاسترجاع في المكتبات

طبيعة نظم الاسترجاع في المكتبات :

على الرغم من صعوبة تعريف نظام استرجاع المعلومات كما يقول (لانكستر) نظراً لاختلاف التطبيقات والنتائج المترتبة على الاسترجاع سواء كانت معلومات نصية أو إشارات ببليوجرافية ، فإن نظام الاسترجاع المقصود هنا : هو النظام الذي لا يزيد الحالة المعرفية للمستفيد من النظام في موضوع معين ، وإنما يحيطه علماً بوجود معلومات عن الوثائق التي تتعلق بموضوع الاستفسار الذي يطرحه الباحث أو المستفيد ، ولذا فإن المعلومات التي يتعامل معها نظام الاسترجاع هي بيانات أو معلومات وسيطة بين المستفيد والمعلومات النهائية التي يسعى إليها الباحثون في المكتبات ومراكز المعلومات .

أما الاسترجاع فيعني الوصول إلى المعلومات المستهدفة عبر سلسلة من الخطوط أثناء البحث في نظام المعلومات الببليوجرافية الذي تكون اللغة العربية فيه هي لغة المعالجة في الإدخال والإخراج عبر المداخل أو الكشافات التي تؤلف نظام المعلومات المستخدم في المكتبات ، أما اللغة العربية فهي الكلمات والمفاهيم التي تعبر عن المعلومات سواء كانت اللغة طبيعية أو اصطناعية مكتوبة بحروف وألفاظ عربية ، ويقصد بالمكتبات جميع أنواع المكتبات ومراكز المعلومات أو ما شابه ذلك من وحدات أو مؤسسات معلوماتية إذا كانت الأنظمة فيها موجهة نحو التعامل مع المفاهيم والموضوعات كمداخل وسيطة بين المستفيد والمعلومات النهائية .

ولقد ظهر خلال السنوات القليلة الماضية العديد من البرامج وقواعد البيانات والنظم المتكاملة في اللغة العربية ، كما قد طرحت من خلال تلك التطبيقات العديد من الحلول والأساليب لمعالجة مشكلات الاسترجاع باللغة العربية سواء كانت تلك النظم للتطبيقات العامة أو كانت موجهة لأغراض استرجاعية خاصة أو محددة بنوعية معينة من النصوص أو المعلومات . ورغم قوة بعض الأنظمة المتاحة وخصائصها الكثيرة والنجاحات المحققة في التعامل مع اللغة العربية إلا أن الحلول المطروحة إما أنها جزئية أو قد تكون على هيئة حزم برامجية عامة تحتاج إلى كثير من التطويع أو التطوير حتى تفي بأغراض البحث والاسترجاع في بيئة المكتبات ومراكز المعلومات .

وإذا كانت قضية معالجة المعلومات باللغة العربية قضية عامة تشغل بال كافة المعنيين بمعالجة البيانات والمستفيدين من المعلومات في كل التخصصات فإن لقضية معالجة المعلومات واسترجاعها فى مجالات المكتبات أهمية خاصة تميزها عن الكثير من مجالات التطبيق الأخرى للأسباب التالية :

١ - إن بيئة الاسترجاع في المكتبات ومراكز المعلومات بيئة مفتوحة على جمهور عريض من المستفيدين وعلى وسائط كثيرة ومتزايدة من وسائط تخزين المعلومات .

٢ - إن نظم التخزين والاسترجاع تعتمد على مواصفات كثيرة موحدة يتم الالتزام بها بحكم أن المكتبة مؤسسة مفتوحة للتعاون وتبادل المعلومات والمشاركة فيها مع المكتبات ومراكز المعلومات الأخرى .

٣ - إن نظم المكتبات تتعامل مع نصوص ومعلومات متزايدة باطراد سواء كانت المعلومات تنمو من داخل المكتبة عن طريق التزويد والفهرسة والتكشيف أو كانت المعلومات ترد من خارج المكتبة على وسائط جاهزة.

٤ - إن لغة الاسترجاع في نظم المكتبات تؤدي دوراً حيوياً لا يضاهيه إلا القليل من المجالات والتطبيقات الخارجة عن علوم المكتبات والمعلومات حين يكون للمفاهيم والموضوعات مداخل مهمة في نظام الاسترجاع ، ولكن أغلب مجالات تطبيق الحاسوب في التخزين والاسترجاع موجهة نحو التعامل مع نصوص محددة أو ثابتة الحقول أو تعتمد على قيم رقمية أو مفاهيم ثابتة . أما البحث والاسترجاع في بيئة المكتبات فعماده اللغات الطبيعية أو الاصطناعية .

٥ - إن للغة العربية بخلاف اللغات اللاتينية خصائص دلالية وتركيبة معقدة تؤثر على فعالية ودقة الاسترجاع بشكل ملح يتجاوز احتياجات الكثير من النظم والتطبيقات الأخرى .

ولذا يجب أن تكون نظم المعلومات مجهزة بالبرامجيات والحلول المناسبة لاسترجاع المعلومات باللغة العربية في مجالات المكتبات المتنوعة .

وحيث إن محور اهتمام الموثقين والعاملين في المكتبات ينصب على اللغة الناقلة للمعلومات فإن قدرة الحاسوب على التعامل مع اللغة هي المحك الحقيقي لنجاح النظم والأساليب المطبقة لمعالجة استرجاع المعلومات عبر مداخلها الثابتة والمتغيرة سواء كان البحث يجري في كشافات العناوين أو الموضوعات أو المستخلصات أو المؤلفين ، أم في النصوص الكاملة المختزنة لأغراض البحث والاسترجاع بواسطة المفاهيم والمصطلحات .

وإذا كان الحاسوب مهيأً أكثر للتعامل مع الأرقام والرموز الثابتة ، فإن لغويات المعلومات ليست ثابتة في معظم الأحيان لأنها تعتمد على الأفكار والمفاهيم وعلى السياقات اللغوية والعلاقات المنطقية التي يحددها الإنسان سابقًا أو لاحقًا ، ويقول (نبيل علي) في معرض حديثه عن اللغة العربية وتقنية المعلومات :

"مصدر الإشكالية إذن ، في كيفية التصدي لمرونة اللغة واتساع مفرداتها ، وتعدد معانيها ، وتباين استخداماتها ، وفي كيفية التعامل مع لبسها ، وغموضها وحذفها وتفشي مجازها ، والكشف عن مضامين كناياتها وبنية جملها ونصوصها " .

والخلاصة أن لقضية الاسترجاع في مجالات المكتبات شقين ، أحدهما يتعلق بالتقنية والبرامجيات أو النظم المتاحة وميزاتها المتنوعة ، والشق الآخر يتعلق بمعالجة اللغة ودلالات الكلمات المستخدمة في تخزين واسترجاع المعلومات . ورغم أن الجانبين متلازمان لمعالجة مشكلة الاسترجاع في النظم العربية إلا أن الحلول المطروحة في إطار الشق الأول يجب أن تكون مبنية على توجهات واحتياجات العاملين في مجالات تنظيم المعلومات واسترجاعها وهم المستفيدون والموثقون والعاملون في مجالات المكتبات ومراكز المعلومات بمفهومها الواسع .

وفي هذا البحث سيتم التطرق إلى العديد من قضايا المعلومات ومناقشة مشكلة الاسترجاع باللغة العربية من زوايا معلوماتية ودلالية مع تلمس الحلول والبدائل الممكنة بالاستناد على أمثلة تطبيقية باللغة العربية . والهدف من ذلك أن نسعى إلى المساهمة في المحاولات الجارية لتصميم وتطوير نظم عربية مرنة ومتجاوبة مع احتياجات المستفيدين في المكتبات ومراكز المعلومات. ومن ثم فاني آمل أن تكون التساؤلات والمعطيات التي يطرحها هذا الكتاب امتداداً لدراسات أخرى أشمل موجهة نحو هذا الجانب الحيوي من علم المكتبات والمعلومات باللغة العربية .

متطلبات تصميم نظام الاسترجاع :

بما أن تصميم قواعد البيانات وعناصر النظم الببليوجرافية ومكوناتها أو خصائصها الآلية قد لا تتغير باختلاف اللغات فسوف ينصب الاهتمام على خصوصيات الاسترجاع باللغة العربية ودراسة التأثيرات المتبادلة بين اللغة والاسترجاع . وسيتم التركيز على لغويات المعلومات وعلى المداخل الموضوعية وطرق عرض المعلومات وتجهيزها حسب احتياجات العاملين في المكتبات العربية وما يصاحب ذلك من مشكلات وحلول . ولن يتم النظر إلى المسائل الإدارية والتنفيذية التي تنطبق على كافة مشاريع التشغيل الآليه في المكتبات وغيرها .

وعند التفكير في تصميم نظام الاسترجاع والمشاركة في تصميم الكشافات المكونة لنظام الاسترجاع العربي يجب أن يكون لدى المكتبي أو محلل نظم المكتبات بعض الأسس والتصورات التي تغطي مختلف جوانب نظام المعلومات في المكتبة وذلك بما يشمل الجوانب التالية :

١ - مشكلات الاسترجاع :

نظم المكتبات من أعقد نظم الاسترجاع لأنها تعتمد على معالجة كم هائل من النصوص التي يتداخل فيها عوامل دلالية وتركيبية تؤثر على صحة المعلومات المطلوب الوصول إليها . وهناك مشكلات استرجاعية تصاحب استخدام الحاسوب باللغة العربية سواء كانت اللغة المستخدمة في الاسترجاع لغة طبيعية أو اصطناعية ، وهذه المشكلات إما أن تكون متأصلة فى اللغة أو موجودة في الحاسوب وفي نظام الاسترجاع . وتحديد المشكلات وحلها يجب أن

ينطلق من مبدأ تسهيل وتبسيط الاسترجاع للمستفيدين مـن النظام بأقصى ما يمكن من الكفاءة والمميزات المتعددة وعدم قبول البرامج المتاحة على علاتها وعدم التنازل عن الوظائف والخصائص التي يتيحها الحاسوب لحل مشكلات الاسترجاع باللغة العربية ، ومدى تجاوب النظام لحل مشكلات الترتيب ومعالجة (أل) التعريف ونحو ذلك .

٢ - نقاط الوصول والمداخل :

معرفة خصائص المداخل المتاحة للوصول إلى المعلومات من عناوين وأسماء وكشافات ومستخلصات والفروق بينها ودورها في الاسترجاع . وتمييز وظيفة البحث في العناوين سواء كان ذلك لأغراض البحث الموضوعي أم لأغراض التزويد والاسترجاع الفوري لعناوين محددة ، ثم معرفة طبيعة اللغات الاصطناعية المستخدمة ومستوى التحليل الموضوعي ، وإمكانات البحث سواء كانت اللغة المستخدمة سابقة الربط أم لاحقة الربط باستخدام روابط الصلة المنطقية للكلمات الدالة ، ويجب النظر إلى أن المداخل المختلفة تقوم بوظائف متمايزة لمختلف الأغراض الاسترجاعية .

٣ - كشافات التباديل :

معرفة الوظيفة المتميزة للتباديل في كافة المداخل والكشافات والتفريق بين الاسترجاع بالسياق التقليدي وبين التباديل باستخدام نقاط الوصول المتاحة في نظام الاسترجاع . هذا إلى جانب معرفة أسس ومكونات التباديل وطريقة عرض الكلمات المفتاحية للمفاهيم المفردة والمركبة ضمن سياقاتها في الكشافات ، وكذلك معرفة المواصفات الخاصة بالتبويب الألفبائي وعرض المداخل وتجزئة

العنوان وقلب الكلمات وطول حقل المدخل وعدد الكلمات القابلة للتباديل في ثنايا النص للعناوين والمستخلصات والمؤلفين .

٤ - لغويات المعلومات :

١/٤ تحديد أغراض قوائم التوقف والمداخل المعينة لاستخدامها مع معرفة المشكلات اللغوية والاسترجاعية والمحاذير المحتملة عند استخدام كلمات التوقف وحلولها المنطقية التي تناسب طبيعة المعلومات والنصوص التي تتعامل معها المكتبة . هذا إلى جانب معرفة الخصائص وطبيعة كلمات التوقف والمنهج المتبع لحصرها وتحديثها .

٢/٤ تعرُّف أسس وقواعد بناء اللغات الاصطناعية من رؤوس الموضوعات والاستفادة منها في تطويع اللغة الطبيعية وإعداد خطط البحوث الآلية.

٣/٤ تعرُّف الأسس اللغوية الخاصة بالتراكيب النحوية ودلالات المفردات والأسماء والحروف وتصاريفها في السياقات المستخدمة في نظام الاسترجاع.

٤/٤ التعرف على طبيعة الكلمات المفتاحية ومستوى خصوصية المصطلحات ومدى تقنينها لدى الباحثين ، وفي الإنتاج الفكري الذي يهم المكتبة .

فكلما كانت اللغة مخصصة ومقننة زاد التحقيق في الاسترجاع سواء كانت اللغة طبيعية أو اصطناعية وبحسب طبيعة الموضوعات . أما إذا كانت اللغة

الطبيعية مرنة وغنية بالمفردات والتراكيب المتفاوتة بسبب غياب ضبط المصطلحات ، فإن اللغة الاصطناعية المحكمة والمخصصة تصبح أكثر توثيقاً، خصوصًا إذا كانت مواكبة لحركة التأليف والتعريب بشكل فوري .

٥/٤ تعرُّف طبيعة المفردات والمشكلات الدلالية الخاصة بالمشترك اللفظي والمترادفات ومدى تأثيرها على الاسترجاع وطريقة معالجتها في السياق للمفردات الدالة وغير الدالة .

٦/٤ معرفة خصائص ومشكلات المفاهيم المركبة وأسلوب ربطها وطريقة معالجتها في مرحلة المدخلات والمخرجات .

٧/٤ تعرُّف الأساليب والحلول المناسبة لمعالجة المفردات غير الدالة واللواصق التي تسبق أو تلحق الكلمات وتأثيراتها على التركيب والاسترجاع ومدى ملاءمة الحلول للتطبيق على نظام الاسترجاع في المكتبات .

٥ - المستفيدون من نظام الاسترجاع :

معرفة طبيعة احتياجات الباحثين والعاملين الذين سيخدمهم النظام بما يشمل وظائفهم وتخصصاتهم ومهاراتهم وسلوكياتهم في البحث والاسترجاع والمستويات البحثية التي يحتاجونها ومدى استجابة النظام لمختلف الاحتياجات المعلوماتية في مختلف أنواع المكتبات ومراكز المعلومات .

٦- طبيعة المعلومات :

معرفة التخصصات الموضوعية الغالبة على مقتنيات المكتبة ونوعية

المعلومات المطلوبة أو طبيعة النصوص المختزنة وحجمها ، وكذلك أشكال الوثائق والأوعية التي تضمها قاعدة البيانات في المكتبة سواء كانت كتباً أو مقالات أو تقارير أو بحوثاً أو غير ذلك ومدى الاتساق في تحليلها ومعالجتها باستخدام اللغة الاصطناعية المتاحة .

٧ - مخرجات نظام الاسترجاع :

معرفة التعامل مع لغة البحث والاسترجاع وبالذات اللغة الطبيعية عند الاستفسار باستخدام روابط الصلة المنطقية والمحددات المتاحة في مرحلة المخرجات ، وكذلك طريقة عرض العناوين ونتائج البحث عند وجود مضاهاة كاملة . ولا شك أن جميع المعارف والمهارات المكتسبة حول نظام المعلومات ومكوناته وخصائصه يمكن قياس فاعليتها بمحك أداء نظام الاسترجاع والوصول إلى المعلومات وتجهيزها بشكل جيد .

٨ - الأجهزة والأنظمة :

عدا المعرفة التقنية وقدرات ومميزات أجهزة التخزين والعرض المتاحة ، يجب أن تكون لدى المكتبي معرفة أكيدة بمواصفات المحارف العربية وبجميع الرموز المستخدمة لتطويع النصوص والتعامل مع الكشافات أثناء إعدادها واسترجاعها على المطاريف . وهناك بعض أجهزة إدخال البيانات التي تتعامل مع الحروف العربية بطريقة تسمح بتمييز أشكال الحروف في نهاية الكلمة مثل (ح، ع ، ي) ونحوها باستخدام شكل واحد دون حاجة لاستخدام المفتاح المحدد للحروف النهائية على لوحة المطراف . كما أن هناك بعض الرموز التي يتم

برمجتها لأغراض التباديل وكلمات التوقف وللبتر الأمامي والخلفي ونحو ذلك من رموز تستخدم لأغراض تمييز أو استبعاد الكلمات . كما يجب أن تكون لوحة المفاتيح حاوية كافة المحارف والرموز المستخدمة في الفهرسة .

وهناك العديد من المميزات التي تتيحها نظم الاسترجاع فيما يتعلق بتسهيلات البحوث والوصول إلى المعلومات وفق مستويات متفاوتة من العلاقات والعرض والتجهيز حسب قدرات الأنظمة ومتطلبات المستفيدين . ومن التسهيلات المهمة في نظم الاسترجاع الروابط المنطقية ودمج المفاهيم على أوجه متفاوتة في التعضيد والتركيب إلى جانب المحددات الزمنية والمكانية والشكلية والوعائية ، وعلاقات المجاورة والبتر الأمامي والوسطي والخلفي ، إلى جانب أساليب الترتيب الألفبائية والزمنية والمنطقية والترتيب المتداخل حسب نقاط الوصول المطلوبة . كما أن بعض النظم تتيح إمكانية البحث النصي في كافة الملفات والحقول بخطة بحث واحدة يتم تخزينها ثم تمريرها على الملفات واحداً إثر الآخر أو دفعة واحدة ، وإجمالا ، فإن نظام الاسترجاع الفعال هو الذي يتيح استعراض الكشافات والاختيار من القوائم بطريقة ميسرة وباستخدام مصطلحات البحث مرة واحدة ومعالجتها دون حساسية للتفاوت المعتاد في إملاء الكلمات وإدخالها أثناء البحث والاسترجاع .

أدبيات الموضوع :

إن تمثيل المعرفة الإنسانية بالحاسوب مرتبط ارتباطًا وثيقًا بتحليل اللغات الإنسانية وتركيبها (الوعر) . وهناك ارتباط قوي بين اللغة والحاسوب ، فتارة يكون الحاسوب هو أداة التحليل اللغوي وتارة تكون اللغة هي وسيلة الحاسوب في تطوير تصميمه ونظمه ، كما يقول (نبيل علي) في دراسته الاستعراضية المستفيضة حول مجمل القضايا اللغوية التي تتعلق باستخدام اللغة العربية في الحاسوب .

ولقد تنامت في السنوات الأخيرة البحوث والدراسات العربية الموجهة نحو تقنية المعلومات ولغويات المعلومات أو اللغويات الحواسبية . ويكفي لإدراك أهمية وتشعب البحوث النظرية والتطبيقية في هذا الخصب الاطلاع على سلسلة وقائع (ندوة استخدام اللغة العربية في تقنية المعلومات) أو بحوث مؤتمرات الحاسب الآلي التي تعقد في الجامعات السعودية أو في غيرها من البلاد العربية ، وهنا يجب أن نشير إلى أن التجارب والأساليب أو الحلول النظرية والتطبيقية المطروحة لمعالجة استرجاع المعلومات باللغة العربية قضية لغوية وحواسبية تتجاوز أبعادها مجالات المكتبات والمعلومات التي نحن بصددها إلى كافة نظم الاسترجاع لمختلف الأغراض . فالقضية التي نتناول جزئية منها في هذا المبحث ذات جوانب متعددة تشكل نسيجا متداخلا من علم المعلومات والحاسوب واللغويات والرياضيات وعلم المكتبات وغيرها من التخصصات . غير أن ما يهمنا في هذا الصدد ما له علاقة مباشرة بمعالجة واسترجاع المعلومات الموضوعية باللغة العربية حسب الأساليب والقواعد المتبعة في مجالات التوثيق والمكتبات .

ويمكن تصنيف أدبيات الموضوع التي تتناول دلالات وتركيب اللغة العربية وتأثيرها على الاسترجاع في المكتبات بما يلي :

أولاً : النصوص :

ويشمل ذلك البحوث والدراسات والإحصائيات التحليلية التي تعالج استخدامات المفردات والتراكيب في المعاجم أو النصوص . ومن البحوث الكثيرة في هذا الجانب بحث الخولي : التراكيب الشائعة في اللغة العربية : دراسة إحصائية .

وقد طبق الخولي دراستة على (٤٤٠٠) عينة من الكلمات التي وقعت في ثمانية وثمانين نصًا مختلفة الموضوعات . وقد جاءت معطيات البحث بنتائج كثيرة عن استخدامات التراكيب اللغوية يهمنا ما له علاقة بالنصوص الحرة المستخدمة في المستخلصات أو في عناوين الكتب . فقد دل البحث على أن أكثر من نصف الكلمات التي نستخدمها هي أسماء تليها في الترتيب الحروف ثم الأفعال .

أما من حيث شيوع استخدامات الحروف فإن واو العطف وحروف الجر أكثر من غيرها ، حيث ترد في ترتيب الاستخدام كما يلي : (في ، الباء ، من ، على ، إلى ، عن) . أما واو العطف فتأتي في رأس قائمة حروف الجر بفارق كبير جدا بينها وبين حروف العطف الأخرى ، أما أكثر التوابع استخدامًا في النصوص العربية فهي الصفات ثم المعطوف (كما بينت ذلك دراسة (الخولي) .

على أن دراسة المفردات الشائعة من الحقول المطروقة كثيرا في علم اللغة. ومن هذه الدراسات المفيدة في معرفة الكلمات الشائعة بحث (داود عطية

عبده) (المفردات الشائعة في اللغة العربية) حيث طبق المؤلف دراسته الإحصائية على (٣٠٢٥) كلمة اختارها المؤلف من دراسات إحصائية أخرى كلها تدور حول المفردات الشائعة في النصوص العربية ، ثم استخلص المؤلف نتائج جديدة عن المفردات الأكثر استخدامًا في النصوص العربية ، وقد اخترت بالترتيب أكثر الكلمات من ضمن الخمسين كلمة الأكثر شيوعًا في الاستخدام ، مع ملاحظة أنني قد أغفلت الكلمات الدالة لعدم أهميتها في هذا المجال كما في القائمة التالية :

في ، من ، على ، إن ، إلى ، كان ، هذا ، هذه ، الذي ، التي ، لا ، ما ، قال ، قد ، عن ، ذلك ، تلك ، كل ، كم ، ثم ، هؤلاء ، إذا ، بين ، يوم ، مع، هى ، بعد ، أو ، حتى ، لكن ، كثر ، غير ، ماذا ، لماذا ، يا ، إلا ، خذ، بعض ، كبير ، أول ، شيء ، عند ، جميع .

هذا إلى جانب كلمات أخرى كثيرة الورود في النصوص من الكلمات غير الدالة . وقد لاحظت أن الكلمات الدالة تقع في مستوى أدنى من حيث تكرار الاستخدام ، كما أوضحت ذلك الجداول التي يضمها بحث (عبده) ، وهذا مؤشر قوي على أن المفردات غير الدالة أكثر شيوعًا في النصوص من الكلمات الدلالية المهمة في الاسترجاع .

وعلى الرغم من أن هذين البحثين طبقا على النصوص المكتوبة المأخوذة من متون عديدة إلا أن لهما أهمية قوية في مجال استرجاع المعلومات ، وذلك نظرا لاحتوائهما على الكثير من المؤشرات الإحصائية حول تصاريف الكلمات واستخدامات المفردات والحروف والتراكيب مما يفيد في دراسة وتحليل المفردات

الدالة وغير الدالة في نصوص المستخلصات والعناوين المستخدمة في الاسترجاع ، ومن ثم يمهد الطريق لدراسة لغويات المعلومات باللغة العربية .

ثانيًا : المعلومات :

إن الكلمات الموضوعية والمفاهيم الدالة هي اللغة الأساسية التي يتم التعامل معها في مجال تنظيم المعلومات . ومع ازدياد أهمية البحوث والاسترجاع بالمدخل الموضوعي تنامت الدراسات الموجهة نحو لغويات المعلومات المستخدمة في المكتبات ومراكز المعلومات وما يتبعها من نظم استرجاع . ويقول (شعبان خليفة) : " بأن البحث في موضوع معين هو بغية رواد المكتبة في الوقت الراهن ويمثل ٩٠٪ من خدمات أي مكتبة أو مركز معلومات " .

وقد أشار Spark إلى العلاقة القوية بين اللغويات وعلم المعلومات في كتابه الدراسي المخصص لهذا الموضوع ، حيث قال ما خلاصته : إن العلاقة القوية بين علم المعلومات واللغة تتضح بما بين التكشيف واسترجاع المعلومات من ترابط يستلزم من المكتبيين معرفة دلالة الرسالة التي تحملها الوثيقة وبالأساليب اللغوية التي يطبقها المكتبيون في أعمالهم التوثيقية .

وقد انصبت لغويات المعلومات المبكرة ضمن عناية المكتبيين العرب بتطوير قوائم رؤوس الموضوعات وإعداد قواعد صياغة المداخل باللغة العربية ، وذلك ضمن اهتمامات (الخازندار ، عبدالهادي ، شعبان خليفة) وغيرهم من الأفراد والمؤسسات الذين تولوا إصدار قوائم رؤوس الموضوعات والمكانز العربية .

ومنذ صدور هذا البحث في طبعته الأولى ظهر في الأدبيات المنشورة حول الموضوع الكثير من الدراسات والمقالات المهمة . ومن الدراسات التطبيقية الحديثة عن دلالات العناوين العربية وأهمية استخدامها في كشاف التباديل

البحث الذي أجراه الباحثان (بشير) و (بكستون) لفحص مدى تمثيل عناوين الدوريات العربية في خمسة حقول هي : الزراعة ، وعلم المكتبات والمعلومات، واللغويات ، والفلسفة ، وعلم الاجتماع ، وذلك بهدف معرفة إمكانية الاعتماد على العناوين في تكشيف الكلمات . وقد استخدم الباحثان منهج مقارنة الكلمات الاسمية الدالة في العنوان مع الجملة المؤشرة إلى الموضوع والتى تقع عادة في مقدمة أو ثنايا المقالات مثل عبارات (يهدف هذا البحث إلى ... ، وهذه المقالة تدرس ... ، وفي هذه الدراسة سوف يحقق المؤلف في ... ، والورقة التالية تعالج ...) ونحو ذلك من عبارات يستخدمها المؤلفون للنص على مضامين مقالاتهم وبحوثهم .

ورغم وجود بعض المفارقات بين دلالات العناوين ومحتوى المقالات في بعض الحقول مثل علم المكتبات والمعلومات والفلسفة فقد وجد الباحثان أن ٧٠٪ - ٨٨٪ من الكلمات الاسمية الدالة في العناوين تماثل الكلمات الاسمية في الجمل المؤشرة على الموضوعات في صلب المقالات

والخلاصة : أنه يبدو أن العناوين العربية تمثل موضوعات المقالات بدرجة جيدة .

ومن الدراسات الحديثة التي تناولت مشكلة الاسترجاع باللغة العربية البحث التجريبي الذي أعده (السويدان) لمعرفة مدى كفاية كلمات العنوان للدلالة على الموضوعات في ثمان وعشرين دورية من تخصصات مختلفة ، وقد تشابهت نتائج البحث مع معطيات الدراسة التي نحن بصددها .

وفي بحث تطبيقي آخر أعده (السريع) تناول الباحث العديد من الصعوبات الفنية التي واجهت المتخصصين عند تصميم نظام (ابن النديم) المستخدم في مكتبات معهد الإدارة العامة حيث تطرق البحث إلى استخدام اللغة العربية

والمشكلات التي صاحبت ترتيب عناصر أسماء المؤلفين وتباديل الكلمات في كشافي المؤلفين والعناوين وكلمات التوقف مع ذكر الحلول التي تم التوصل إليها لمعالجة المشكلات والقواعد التي تحدد كيفية إدخال واسترجاع المعلومات .

كما عرض (البخيت) لبعض العوامل اللغوية المؤثرة على دقة الاسترجاع بالاعتماد على العناوين في قواعد البيانات العربية مثل (أل) التعريف وحروف العطف والجر وكذلك بعض أساليب البتر واستخدامها في البحث .

كما ناقش (الأطرم) مشكلة استخدام اللغة الطبيعية في استرجاع المعلومات باللغة العربية . وأشار في مقاله إلى أبرز العقبات الدلالية والآلية التي يجب التغلب عليها من أجل تصميم أنظمة معلومات تناسب طبيعة اللغة العربية ، وذلك ضمن مشروع بحث متكامل تشرف عليه مدينة الملك عبدالعزيز للعلوم والتقنية .

أما (حشمت قاسم) فتناول في مقاله المبكر الخلفية التاريخية لكشافات الكلمات المفتاحية ، ثم أشار إلى بعض الجوانب النظرية والتطبيقية لاستخدام اللغة العربية في كشاف الكلمات الدالة في السياق ، كما تطرق (الهبائلي) إلى العلاقة القوية بين المعلومات واللغويات حيث أشار بشكل عام إلى الأسس النظرية لدلالات اللغة دون أن يركز على مشكلات الدلالة والتركيب في اللغة العربية .

ولا شك أن دراسة لغويات المعلومات العربية تستند على خلفية غنية من الأدبيات المبكرة حول أسس وصياغة اللغات الاصطناعية . وما هذا الكتاب إلا محاولة لتلخيص ومعالجة أبرز المشكلات الدلالية والتركيبية للغات الطبيعية المستخدمة في الاسترجاع في المكتبات العربية .

الفصل الثاني

لغويات المعلومات

- النص
- المفهوم
- أشكال المفاهيم
- المصطلحات الأجنبية
- اللغة الطبيعية والاصطناعية
- خصائص الاسترجاع باللغة الطبيعية
- مشكلات الاسترجاع باللغة الطبيعية



الفصل الثاني
لغويات المعلومات

اللغة نظام من الإشارات التي تعبر عن الأفكار كما يقول (سوسور)* وتنبثق العلاقة بين المعلومات واللغة من أسس علم اللغة الألسنية، وهي علم الدلالة (الصرف) والتركيب (النحو) . فعلم الدلالة يتناول دراسة معاني الكلمات والألفاظ ، أما علم النحو والتركيب فيتناول ترتيب عناصر الجملة والعلاقات بين المفردات المستخدمة في وصف المعلومات وتناقلها عبر الوسائط المختلفة. وفي مجال المكتبات تستخدم الإشارات الرقمية والحرفية أو النصوص المسجلة التي تعبر عن المفاهيم التي يمكن معالجتها في عمليات التحليل الموضوعي والاسترجاع .

أما المعلومات فهي ضد المجهولات ، أو كما يقول القاضي أبو بكر الباقلاني في الإنصاف : بأن المعلومات من المدركات والمبصرات والمسموعات، أو هي المفاهيم الدالة على كيانات وعمليات أو مجردات مما يأتي على هيئة بيانات حقائقية أو تفسيرية تتألف منها المعرفة البشرية .

النص :

النص مجموعة من الكلمات والحروف أو الجمل والعبارات التي تشكل بنية دلالية سواء كانت نابعة من النص أو من خارجه وسواء دلت على مفاهيم ثابتة أو متغيرة ، والنص كيان معرفي له دلالات إيحائية أو مفاهيم قابلة للتفسير

* دي سوسورا أو سوسير كما في بعض المصادر.

المتغير حسب وجهات نظر المتلقي والمجالات الموضوعية أو مناهج وأغراض الدراسات المختلفة سواء كانت دينية أو لغوية أو أدبية أو قانونية أو سياسية أو نفسية وغير ذلك من حقول المعرفة .

وتتفاوت لغة النص حسب وظيفته وأسلوب الكتابة الذي يؤدي دوراً مهمًا في تشكيل النص وإيصال المعلومات الحقائقية أو التفسيرية .

ويمكن تصنيف النص من حيث البناء اللغوي إلى ثلاث فئات من الكلمات حسب الصفات الدلالية :

١ - المفهوم الدال ، وهي المصطلحات والكلمات التي تحمل دلالات موضوعية مستقرة ومتفقًا عليها لتسمية المفاهيم أو الأشياء والأفكار.

٢ - المفهوم غير الدال ، وهي المعاني العامة أو النسبية المتغيرة والتي تلازم المفردات التي لا يمكن تحديد دلالالتها الموضوعية بشكل ثابت .

٣ - الحروف والأدوات النحوية الملازمة لتكوين النص باللغة الطبيعية .

ويدخل في إطار النصوص التي يتم التعامل معها في المكتبات : القرآن الكريم ، والحديث النبوي ، والنصوص الأدبية والإبداعية ، مثل الشعر والحكم والأمثال ، والمقتبسات ، والخطب ، والروايات ؛ وكذلك النصوص القانونية وغيرها من كتابات مهمة .

وعلى مستوى نظم المعلومات فإن النص يختزن ويسترجع لذاته وبحسب الإمكانات والمداخل المتاحة للوصول إلى النص المستهدف . وبما أن النص محور الاسترجاع فإن النص ومكوناته من المفردات والحروف وأنماطها الصرفية وحركاتها الإعرابية قد تصبح ذات أهمية في الاسترجاع .

أما الأساليب المتبعة لتوثيق النصوص واسترجاعها في المكتبات فلا تتجاوز التحليل الشكلي للنص في معظم النظم . ورغم أن كافة النصوص تحمل دلالات وإيحاءات إلا أن الموثقين في التحليل الموضوعي لا يتدخلون في تفسير دلالات الألفاظ وتحديد مفاهيم النص ، وإنما يشيرون إلى الشكل البنائي للنص أو نوعه الأدبي المتعارف عليه ، فتوصف الأعمال الإبداعية مثلاً بأنها نثر ، شعر ، قصيدة ، رواية ، قصة ، مسرحية ، خطبة ، أو أحكام وأمثال . ويترك تفسير النص وتحديد مفاهيمه غير الموضوعية للمحللين أو النقاد والمفسرين حسب مجالات النص .

ويدخل في إطار العمل التوثيقي تكشيف الكلمات الدالة فقط ويتم تطبيقه على عناوين الوثائق من كتب وبحوث ومقالات أو في المستخلصات التي تشكل أحد العناصر المكونة لنظام الاسترجاع في المكتبات .

أما تكشيف النصوص الذي يعتمد على كافة مفردات النص على أنها وحدات للاسترجاع ودون تحليل موضوعي للمفاهيم فيتجاوز أعمال المكتبات إلى التأليف المرجعي الذي يحتاج إلى نظام استرجاع خاص بوظيفة النص وطبيعته واحتياجات المستفيدين منه . ومن هنا نجد أن نظم تكشيف نصوص القرآن تختلف عن نظم تكشيف نصوص الحديث أو غير ذلك من نصوص رغم أن وحدة الاسترجاع واحدة وهي المفردات . ونجد مثلاً أن (معجم الأمثال العربية) تأليف (مراد) ، وهو كشاف نصوص يورد (٢٣) تصريفا متولدة من مادة (ك ث ر) مثل (كثر ، أكثر ، استكثر ، الكثر ، الكثرة ، مكثر ، مكثار، كثير) وغيرها من المفردات المعجمية التي لا تدل على مفاهيم محددة ومستقلة رغم أنها قد وردت في سياقات الأمثال العربية المحصورة .

المفهوم :

المفـهوم هـو المعنى الـدال المـلازم للكلمـات التـي يتـألف منـهـا النص . والكلمـة في المعنى اللغوي هي اللفظ المفرد . أمـا معناها الاصطلاحي في علم المكتبات والمعلومات فهي أصغر وحدة في اللغة تستطيع نقل معنى محدد في ذاتها . وهي قـادرة على البقـاء كوحدة مستقلة في جملة . أما المصطلح فهو كلمة أو عبارة تستعمل للدلالة على مفهوم .

والمفـهوم كمـا تعرفـه (المواصفـات القيـاسيـة العربيـة ٥٧١ - ١٩٨٣) هو «وحدة من وحدات الفكر تستـخدم لبناء المعرفة عن العـالم المحيط وإدراكه». أما معجم جمعيـة المكتبات الأمريكية فيعرف المفهوم بأنه «إدراك نمط القيم والبنى الذي يمكن العـقل من تجسـيـد الأشيـاء من خلال مـسـمـيـات مـتـواترة وثابتة».

وتحتاج المفاهيم في عمليـة الاتصال إلى لغة واضحة أو كلمات موحدة بين المرسل والمتلقي حتى يكون بالإمكان تخزينهـا ومعـالجتهـا واسترجاعـهـا ، ثم إدراكها على أنها معلومات أو بالأحرى إشارات دالة ومداخل موضوعية . وبما أن المفهوم ملازم للنص فإن الموثقين يفرزون النص وينتقون الكلمات الموضوعية الدالة على موجودات أو أشياء أو نشاطات أو قيم مجردة أو خواص وصفات ملازمة لمفاهيم أخرى ثابتة في لغة المعلومات . والدلالة لا تتعلق بالمعلومات ونوعيتها وإنما باللغة المستخدمة لوصف المعلومات .

وتعتمد المكتبات في استرجاع المعلومات على المنهج المفـهومي سواء كـانت اللغة اصطناعيـة أو طبيعيـة . وفي حالة الاعتـمـاد على اللغة الطبيعـيـة في الوصول إلى المعلومات فإن المداخل أو الكشافات قد تتضمن العديد من الحروف

والمفردات التي تشكل بمجملها نصًا يحتوي على كلمات دالة وأخرى غير دالة. والألفاظ غير الدالة ليس لها أهمية في تكوين العناصر الموضوعية التي تفيد في الاسترجاع .

ولذا فإن المفردات غير الدالة تعيق الفهم والوصول السريع إلى المعلومات أثناء البحث الموضوعي . أما تقييم أهمية المفاهيم فيتم بالنظر إليها من زاويتين : الزاوية اللغوية ، والزاوية المعلوماتية أو احتياجات المستفيدين حسب قواعد وسياسة التكشيف المتبعة .

لذا لا بد من التمييز بين النص الذى يتم تخزينه ومعالجته داخل نظام مغلق وموجه للتعامل مع المفردات داخل سياقاتها الضيقة أثناء الاسترجاع ، وبين المفاهيم التي لها دلالات ثابتة نسبيًا ومستقلة ضمن الإطار العام لنظام المعلومات الموجه لفئات عريضة من الباحثين والمستفيدين في المكتبات ومراكز المعلومات .

أشكال المفاهيم :

تأتي المفاهيم في أشكال لغوية مختلفة في اللغة الطبيعية سواء وافقت سمات التركيب الاصطناعي للمصطلحات أم لا . فالمفاهيم قد تكون كلمة واحدة مثل (المرأة) أو مفهومًا مركبًا من كلمتين أو أكثر مثل (تعليم المرأة) . ويتم تركيب المفاهيم في اللغة الاصطناعية وفق القواعد الواجب الالتزام بها لتنظيم المعلومات وليس حسب قواعد اللغة من نحو وصرف . فمثلا :

لا يجوز في اللغة الاصطناعية استخدام الفعل للدلالة على مفهوم كما لا يجوز الربط بين مفهومين مستقلين أو استخدام مفردات غير دالة رغم أن ذلك

جائز في اللغة الطبيعية . ورغم التفاوت في الدلالة والتركيب بين لغات المعلومات إلا أنه يمكن التمييز بين اللغة الطبيعية والاصطناعية في أشكال صياغة المفاهيم كما في الجدول التالي :

اللغة الطبيعية	اللغة الاصطناعية
صيانة الطرق	الطرق - صيانة
دليل التلفون	الهاتف - أدلة
دليل تعليم المرأة في السعودية	المرأة - تعليم - السعودية - أدلة
كيف تحج أيها المسلم	الحج
الشمس والقمر	الشمس
	القمر
أيكولوجية الإدارة	البيئة الإدارية
أوبك	منظمة الأقطار المصدرة للبترول

أما الأشكال الأساسية للمفاهيم فهي : المفاهيم المفردة ، والمفاهيم المركبة أو المعقدة ، ثم الأسماء وما يتفرع من هذه الأشكال من صيغ وتفريعات يمكن استعراضها حسب الأسس العامة لصياغة اللغات الاصطناعية مع بعض الإشارات المقارنة للغة الطبيعية .

المفاهيم المفردة :

يأتي المفهوم ضمن كلمة واحدة ، أو لفظة على هيئة واصفة ، أو مصطلح . سواء ورد ذلك في اللغة الطبيعية أو في اللغات الاصطناعية . فالمفهوم المفرد

هو اللبنة الأساسية في لغة المعلومات ما دام المفهوم ذا دلالة مستقلة مثل : (الإسلام ، الأطباء ، الطب ، الحب ، الطاقة ، المرأة ، مكة ، العرب ، أفريقيا ، المعادن) .

ومع أن قواعد صياغة رؤوس الموضوعات تميل إلى تفضيل الكلمة المفردة في تمثيل المفاهيم بسبب سهولة معالجة المفهوم المفرد وسهولة تذكره واستخدامه في البحث والاسترجاع إلا أن اللفظ المفرد بسبب ضعف خصوصيته فهو أقل استخدامًا في صياغة المفاهيم باللغات الاصطناعية كما يقول (عبدالهادي) .

ورغم وضوح دلالات المفاهيم المفردة خصوصا الأسماء إلا أنه يتم معالجتها في الإنتاج العلمي من زوايا منهجية وتطبيقية مختلفة . لذا فإن الكلمة المفردة قد لا تفي بأغراض البحث والاسترجاع بسبب ارتفاع نسبة الاستدعاء على حساب دقة التحكم بالمعلومات . ومع أن اللفظ المفرد يدل على مفهوم عام إلا أن المقيدات المختلفة والتفريعات الوجهية تخصصه في اللغات الاصطناعيه مثل : الطاقة - تكاليف - إحصائيات - فرنسا - ١٩٩٤م. بينما قد يزيد السياق في تخصيص المفهوم المفرد بشكل أكبر اعتمادا على اللغة الطبيعية سواء كانت في العناوين أو المستخلصات . فإذا كان احتمال التركيب المفهومي ثابتًا ومحدداً في اللغات الاصطناعية فإن المفهوم المفرد قد يركب مع مفاهيم أخرى بشكل مطلق في اللغة الطبيعية . وهذا ما يجعل المفهوم المفرد قابلا للتكيف حسب الاحتياجات البحثية للمستفيدين بينما اللغات الاصطناعية مهما بلغت من الحصر ومقدار المقيدات فإنها لا تستطيع مجاراة اللغة الطبيعية في الربط بين مفهومين فأكثر سواء وافق ذلك نسق التركيب الاصطناعي في المقيدات الوجهية والشكلية والمكانية والزمانية أم لا .

فالمقيدات الوجهية مجال خصب تتفوق فيه اللغة الطبيعية على اللغات الاصطناعية . وقد سرد (شعبان خليفة) الكثير من التفريعات والمقيدات التي تضاف إلى رؤوس الموضوعات على سبيل المثال لا الحصر لصعوبة ذلك.

أما من ناحية التركيب فإن المفهوم المفرد قد يركب مع العديد من المفاهيم الأخرى التي تجعله مفهوما مركبًا من الناحية اللغوية باللغة الاصطناعية كما سيأتي . فمثلا نجد المفهوم (معادن) يرد مركبًا باللغات الاصطناعية مثل :

- استخراج المعادن
- علم المعادن
- التنقيب عن المعادن
- صناعة المعادن

وكذلك في اللغة الطبيعية مثل العناوين التالية :

- أساسيات قطع المعادن
- تشكيل المعادن بدون قطع
- تشكيل المعادن بالضغط
- حسابات قطع المعادن
- تآكل المعادن
- قواعد تشغيل المعادن

ودون التركيز على دقة دلالة المفاهيم باللغة الطبيعية مقارنة برؤوس الموضوعات كما في الأمثلة السابقة ، فإن الاسترجاع في حالة الربط اللاحق في الكشافات والمستخلصات يعتمد على المفهوم المفرد سواء جاء جزءاً مركباً مع غيره من اللغات الاصطناعية أو مخصصًا ومركبًا في اللغات الطبيعية سواء

ظهرت الكلمة الدالة في بداية المفهوم الاصطناعي المركب أم جاءت الكلمة تابعة. وهكذا نرى أن لمفهوم (المعادن) في اللغة الطبيعية علاقات مفهومية وتركيبية أخرى تتجاوز مفهوم (الصناعة)إلى مفاهيم أخرى أدق مثل : (القطع، التشكيل ، الضغط ، الحسابات ، التآكل والتشغيل) .

والخلاصة : أن المفهوم المفرد يحتاج في معظم الأحيان إلى تقيد وجهي أو ربط تركيبي بمفهوم آخر سواء كانت لغة الاسترجاع اصطناعية أو طبيعية وسواء كان الربط سابقًا أو لاحقًا . من هنا يقل استخدام المصطلح المفرد دون ربط أو تخصيص أثناء البحث والاسترجاع إلا في حالة الأسماء المفردة للأشخاص والأشياء أو بعض المفاهيم الجزئية في العلوم أو في حالة البحث الاستطلاعي المبدئي . هذا على أساس أن مجالات البحث والاسترجاع تقع ضمن إطار التغطية الموضوعية لأهداف نظام المعلومات من الناحية الكمية والنوعية للمعلومات المختزنة .

المفاهيم المركبة :

المفهوم المركب هو المصطلح أو رأس الموضوع المكون من كلمتين فأكثر . والمفهوم المركب متعدد من الناحية اللغوية ، ولكنه متحد من الناحية الدلالية حيث يدل المفهوم المركب على موضوع واحد مميز . ولذا فهو أدق في تحديد المفاهيم وفي التعبير عن العلاقات المنطقية بين زمرة من المفاهيم المفردة . كما أن المفاهيم المركبة أكثر استخدامًا في صياغة المفاهيم المستخدمة في اللغات الطبيعية والاصطناعية .

ويتم تركيب المفاهيم حسب القواعد النحوية والقواعد الاصطناعية فيتم التعبير عن الموضوعات المعقدة حسب القواعد اللغوية التي تشمل :

١- الصفة والموصوف ، مثل :

- الطاقة الهوائية
- القانون الدولي الخاص

٢ - المضاف والمضاف إليه ، مثل :

- طب الأسنان
- إدارة الفنادق
- رؤوس الموضوعات العربية

٣ - الروابط النحوية : ويشمل ذلك حروف الجر والعطف والظروف الزمانية والمكانية ونحو ذلك ، مثل :

- الإسلام في الهند
- الإدارة بالأهداف
- المرأة كمحامية
- الرقابة على المطبوعات
- العرب قبل الإسلام
- المركزية واللامركزية
- الاتصال عن بعد
- التأمين ضد الحريق

كما يتضاعف التركيب اللغوي في اللغات الاصطناعية انطلاقًا من المفهوم الأساس المفرد أو المركب إلى المقيدات وتفريعات الموضوعات وكذلك قلب الرأس حسب القواعد الاصطناعية المتبعة في صياغة رؤوس الموضوعات مثل :

الأطباء - أدلة

المرأة - تعليم

الطاقة الهوائية - تكاليف - إحصائيات - ألمانيا - ١٩٩٤م .

وبغض النظر عن المقيدات غير الوجهية مثل الشكلية والمكانية فإن أفضل المفاهيم صياغة هي الثنائية الخالية من الروابط . وأقوى التراكيب في اللغة العربية المضاف والمضاف إليه لأن المصطلح في الغالب يتألف من مفهومين مستقلين دلاليًا . أما أضعف القواعد اللغوية لتركيب المفاهيم فهي التي تستدعي استخدام حروف الجر أو العطف ونحوها من الروابط والمفردات غير الدالة خصوصا إذا كان هناك بدائل لغوية أو قواعد اصطناعية كافية لتركيب المفهوم بشكل محكم دلاليا .

ولكن المفاهيم المركبة قد تتجاوز الثلاث كلمات فتتعقد صياغاتها وتشمل العديد من حروف الجر والعطف والظروف الزمانية والمكانية بل والمفردات غير الدالة . أما رؤوس الموضوعات النصية التي تتجاوز الثلاث كلمات في سياق طبيعي فليست من اللغات الاصطناعية ؛ بل من اللغة الطبيعية إذا كانت دالة على مفاهيم متعددة وليست أسماء ، مثل :

الجلاء البريطاني عن مصر

المساعدات الأولية عند المرض والإصابة

اشتراك العاملين في الإدارة والأرباح

الممتلكات العقارية ونظام الضرائب

الممتلكات المقيدة في العلاقات الصناعية

تأمين مسئولية أصحاب العمل

التأمين على الحياة التجارية

النفايات والتخلص منها

رحلات الفضاء إلى القمر

الطاقة القابلة للتجديد

مراجعة حسابات الضرائب

ضريبة الأرباح غير الموزعة

النجاح في الأعمال التجارية

تأهيل القائمين بالاتصال الجماهيري

تصنيف الرسائل الواردة إحصائياً

الترجمة والترجمة الفورية

اقتصاديات ما بين الصناعة

ورغم أن بعض هذه الرؤوس المعقدة لا تتجاوز الثلاث كلمات إلا أن ضعف صياغتها يتجاوز عيوب اللغة الطبيعية وذلك بسبب الترجمة الحرفية الركيكة من مكانز وقوائم رؤوس موضوعات أجنبية . ولعل من أبرز العيوب التركيبية

والدلالية في رؤوس الموضوعات السابقة هي تكرار المفاهيم في الرأس الواحد وكثرة الكلمات غير الدالة والصيغ الإعرابية غير المحكمة ونحو ذلك من مفردات زائدة تخل بتركيب المفهوم .

ولو نظرنا إلى المفاهيم المركبة من الناحية اللغوية والاسترجاعية لوجدنا أن الصفة تابع مخصص للمفهوم الموصوف ، والمضاف يحدد دلالة المضاف إليه . ثم إن حروف الجر والعطف والظروف المكانية والزمانية تخصص المفاهيم المركبة حسب المعاني اللغوية للروابط النحوية . وكلها تعمل بمثابة الربط المسبق للمفاهيم المفردة من أجل استخدامها في وصف المعلومات واسترجاعها .

وبما أن المفهوم المركب يقع ضمن كلمتين فأكثر فإن ذلك يسبب مشكلات في تحديد العنصر الأهم في ترتيب مداخل الكشافات حيث يفضل أن تكون الكلمة الدالة أو المفتاحية في أول المدخل وكذلك الأمر بالنسبة للبحث والاسترجاع إذ قد يستخدم الباحث الجزء التابع والذي لا يظهر في حالات الربط المسبق للمفاهيم المركبة باللغات الاصطناعية . ولعلاج هذه المشكلة يتم اللجوء إلى مبدأ الرأس المقلوب ، مثل :

الطرق ، صيانة

الفنادق ، إدارة

ويقول (شعبان خليفة) إن المضاف والمضاف إليه في اللغة العربية مجال خصب لعملية القلب بينما الصفة والموصوف في اللغة الانجليزية هي المجال المقابل للقلب . كما تتجاذب الصياغة اللغوية المفهوم المركب بين الصفة والموصوف وبين التقسيم الوجهي في اللغات الاصطناعية مثل :

الصحف السعوية

الصحف - السعودية

والواقع أن المصطلحات تتأثر بالترجمة من اللغات الأجنبية إذ نجد أن تركيب بعض المصطلحات يتم حرفيًا مع حروف الجر مثلاً ، بينما التركيب العربي قد لا يستدعي استخدام حروف الجر ، فمثلاً رأس (المرأة كمحامية) قد يقابله في اللغة الطبيعية المرأة المحامية أو المحاميات و (الرقابة على المطبوعات) يصبح (رقابة المطبوعات) كما أن المضاف والمضاف إليه قد يعوض عن استخدام الروابط الموجودة في اللغة الإنجليزية مثل (حرية التعبير) Freedom of speech.

والخلاصة : أنه كلما زادت المفاهيم تركيبا زدات الحاجة إلى النص أو اللغة من أجل التعبير عن المفاهيم المستهدفة بشكل أدق . غير أن استخدام اللغة الطبيعية قد يلغي دور اللغة مع القواعد المتبعة لصياغة رؤوس الموضوعات المركبة . وذلك يشمل الحالات الإعرابية من الصفة والموصوف والمضاف والمضاف إليه والروابط أو حروف الجر والعطف وكذلك قلب رؤوس الموضوعات والمقيدات الوجهية ، ففي حالة الربط اللاحق للمفاهيم المفردة يضعف دور قواعد النحو والتركيب اللغوي كما يقول Auri نقلاً عن Austin فالربط بين المفاهيم بعوامل الصلة المنطقية (و ، أو ، غير) يحل محل النحو والتركيب . ويتم تركيب المفاهيم عند الاسترجاع بغض النظر عن أصل المفاهيم كما يتم التخلص من المفردات والروابط اللغوية غير الدالة . أما المقيدات الشكلية والزمنية ونحوها فيمكن معالجتها آلياً فتصبح على هيئة مسارد يتم ربطها مع المفاهيم الموضوعية الأخرى باستخدام روابط الصلة المنطقية .

وإذا كان الربط اللاحق للمفاهيم باللغة الطبيعية يلغي دور اللغة والتركيب فهل يصدق هذا المبدأ على اللغة العربية أيضا ؟ وهذا ما سنعالجه عند التطرق لمشكلات الاسترجاع باللغة الطبيعية .

الأسماء :

الأسماء أبلغ في الدلالة على المسميات لأن المسمى كيان مادي ثابت لا يتجزأ إلى مفاهيم أخرى مهما بلغت كثافة المعلومات مع أن بالإمكان إضافة مقيدات أو تفريعات كثيرة في الاسم نفسه ، وإذا نظرنا إلى الأسماء باعتبارها موضوعات أو مفاهيم دالة على الأشخاص أو الهيئات أو الحيوانات والجمادات فإن اللغات الاصطناعية قد لا تحصرها بسبب كثرتها وتنوعها واختلاف المجالات الموضوعية للمسميات حسب اختلاف حقول المعرفة . وقد أكدت الدراسات الحديثة تمايز الاحتياجات المعلوماتية للباحثين في الإنسانيات باعتمادها على الاستفسارات التي تنصب على البحث عن أسماء الأشخاص والمسميات الجغرافية والتاريخية Bates .

وتعتمد طريقة الاسترجاع على صياغة الاسم سواء كان مفرداً أو مركبًا مثل:

المعري

أبو الطيب المتنبي

الفيل

وحيد القرن

بحر قزوين

تهامة

سابك (الشركة السعودية للصناعات الأساسية)

وكالة الأنباء السعودية

واسترجاع الأسماء المفردة سهل . أما عند استرجاع الأسماء المركبة فيتم الاعتماد على العنصر الثابت أو المميز في الاسم المركب . وهي في الغالب الأسماء الأخيرة واسم العائلة للأفراد أو الاسم الثابت أو المعياري للدول مثل (السعودية ، مصر ، سوريا) . كما أن بعض الأسماء المركبة قد تستلزم الربط بين عناصر الاسم واستخدام الواو للربط المنطقي من أجل الوصول إلى مضاهاة موفقة خصوصا إذا كانت عناصر الاسم كثيرة التكرار في المدخل.

وإلى جانب المداخل الاستنادية للأسماء في مداخل المؤلفين ورؤوس الموضوعات فإن الأسماء متاحة للاسترجاع باللغة الطبيعية عبر العناوين والمستخلصات . على أن اللغة الطبيعية للعناوين قد تقصر في الإفصاح عن الأسماء رغم أن المعلومات تنصب على السير والمذكرات الشخصية للأفراد . وهذا يعني أنه لا بد من اللجوء إلى المداخل الموضوعية المقننة من أجل الحصول على معلومات عن الأسماء المستهدفة كما في العناوين التالية :

الأيام	طه حسين
أنا	عباس محمود العقاد
البحث عن الذات	أنور السادات
حياتي	أحمد أمين

ورغم أن للأسماء دلالة في التأكيد على وجود معلومات عن الأشخاص أو المسميات الأخرى إلا أن الأسماء تأتي أجزاء من المفاهيم المركبة باللغة الطبيعية أو الاصطناعية دون أن تدل الأسماء على هوية الأفراد أو ذواتهم وإنما يقصد بها مفاهيم موضوعية مختلفة تماماً عن الشخص . ويدخل ضمن ذلك المفاهيم

والنظريات أو المركبات الكيميائية والمقالات والمكتشفات أو المشاريع المنسوبة للأشخاص من العلماء والمخترعين أو السياسيين ونحوهم . كما قد تتركب المفاهيم الموضوعية مع أسماء جغرافية مثل المؤتمرات والمعاهدات والاتفاقيات العسكرية والسياسية والاقتصادية كما في الأمثلة التالية :

وعد بلفور

نظرية ماسلو

طريقة منتسوري في التربية

هندسة إقليدس

ملخصات شوم

مذنب هالي

اتفاقية الطائف

اتفاقية برن

ومع أن الأسماء أقل عرضة للتغييرات التركيبية خصوصاً الأسماء الشخصية إلا أن هناك بعض العوامل التي تؤثر في دقة الاسترجاع باللغة الطبيعية مما يدعو إلى ضرورة الضبط الاستنادي للأسماء في كافة مداخل نظام المعلومات كما يلي :

١ - تماثل الأسماء واختلاف المسميات أو المفاهيم المقصودة .

٢ - تغير أسماء الهيئات أو الأشخاص .

٣ - وجود الأسماء المستعارة للأشخاص مع الأسماء الحقيقية .

٤- وجود مختصرات للأسماء تختلف عن الاسم الكامل .

٥ – عدم وجود عنصر مميز داخل سياق أسماء الهيئات مثل (الشركة العربية السعودية) أو ضمن أسماء الأشخاص الشائعة بكثرة مثل (أحمد علي محمد).

٦ – استعارة مسميات مغايرة واستخدامها في أسماء الهيئات مثل تهامة، الأهرام .

٧ – وجود مفاهيم موضوعية أو نظريات مربوطة بالاسم .

٨ – وجود كلمات توقف غير دالة في ثنايا الاسم المركب مثل (ابن ، أبو) ونحو ذلك .

٩ – مدى تكرار الأسماء المتماثلة للأشخاص المختلفين أو الهيئات المختلفة.

كما تتعرض الأسماء العربية إذا كانت مشتقة أو مبنية على مسميات أخرى لمشكلة ازدواج المعنى بين أسماء الأعلام المفردة وكيانات أو مفاهيم أخرى عند البحث في اللغة الطبيعية . ولأسباب كثرة الأسماء المشتقة من صفات منسوبة للمواقع الجغرافية والقبائل والمذاهب والمهن فإن بعض الأسماء تماثل الصفات المخصصة لمفاهيم أخرى في الموضوعات المركبة مثل :

الأندلسي	الصيني	الجبري
الحربي	الكوفي	الجوهري
الجزائري	المالكي	الشافعي
السعودي	الورقي	الحنبلي
الكويتي	الحديدي	الكردي
المصري	المعماري	الجندي
الموصلي	الوادي	المحامي
اليمني	الوتري	النووي

وتتراوح الدلالة بين الاسم والصفة العامة أو مفاهيم أخرى مغايرة للاسم الشخصي . وأكثر ما تتضح مشكلات ازدواج الدلالة بين أسماء الأعلام والمفاهيم في نظم استرجاع المعلومات التي تتيح إمكانية البحث النصي الحر بالمفردات في كافة الحقول دون تمييز بين الموضوعات والعناوين والأسماء، وهذا الأمر قد يؤدي إلى بعض المشكلات الدلالية لكثير من الأسماء التي يحتمل أن تدل على مفاهيم أو صفات مختلفة كما في الأسماء التالية :

أديب	عمارة	حرب	سلام	كرم	نجم
إلهام	عمر	صديق	سيف	مطر	شرف
جمال	سيد	توفيق	عبيد	الدنيا	شمس
عبادة	ديرة	سحر	سعادة	حجر	قبس
رمزي	جهاد	سليم	زيتون	آمال	قمر
شريف	نضال	ربيع	بدر	رجاء	هلال
خليفة	سلامة	رياض	غزوة	رمضان	تركي
قحطان	عدنان	منى	سهام	إسلام	عربي
الصدر	الصلح	أمثال	نخلة	قادة	صادر
طالب	معروف	ميلاد	سائد	عيد	صقر
خليل	وجيه	خلاف	جبر	الجمل	مها
الزجاج	جواهر	ياقوت	واضح	جودة	مسلم

كما أن بعض الأسماء تأتي في تهجئات مختلفة مثل (إسحاق ، اسحق ، جودت ، جودة ، ياسين ، يس) وفي صيغ نادرة أو غير معتادة مثل الأفعال (يزيد ، يعرب ، يفوت) . كما تأتي الأسماء في صيغ تبدو وكأنها مفردات

غير دالة مثل (واضح ، أفضل ، سائد ، ألطف) ونحوها . هذا إلى جانب الأسماء الأجنبية التي قد تتوافق بعض عناصرها بالصدفة إما مع كلمات دلالية أو مع مفردات غير دالة مما يستلزم التنبيه لها عند تحديد كلمات التوقف أو عند البحث الحر والاسترجاع باللغة الطبيعية مثل :

و . س . (هولاند)

(وين) (هو) سمان

ول ديور (أنت)

جان (شك) جرو

(ليل) م . سبنسر

(فردي) آي أور

(جيل) دولوز

جون (بون)

(أن) ، ايشو

فيليب (ستيل)

لويس (دوس)

د . أ . (كاميرون)

فالكلمات الدالة التي بين قوسين لها مماثلات عربية سواء كانت الكلمات دالة على مفهوم أو غير دالة .

وهكذا نرى أن الأسماء مهمة جدا في البحث الموضوعي سواء كانت مقننة في قواعد الضبط الاستنادي وفي اللغات الاصطناعية أو جاءت ضمن المداخل الأخرى باللغة الطبيعية . وما عدا تقنين مداخل الأسماء في الفهرسة الوصفية

فإن دراسة دلالات الأسماء وتركيبها وتأثير ذلك على الاسترجاع لم يعط حقه من الدراسة والبحث باللغة العربية كما سنرى لاحقاً عند دراسة كشافات الأسماء .

وأخيراً فإن مشكلات دلالة الأسماء سوف تظهر بشكل واضح في قواعد البيانات الضخمة ، وعند استخدام نظم المعلومات التي تتيح الاسترجاع في الحقول المدمجة باللغة الطبيعية .

المصطلحات الأجنبية :

يكثر في النصوص وفي عناوين المقالات والكتب العربية استخدام المصطلحات الأجنبية المكتوبة بحروف عربية . وتتدرج المصطلحات الأجنبية من حيث الاستخدام والشيوع بين المصطلحات العلمية والأكاديمية التي تصف موضوعات واسعة أو مفاهيم ضيقة إلى جانب المصطلحات العامة المستخدمة على نطاق واسع في وصف الأجهزة والآلات . كما تغطي المصطلحات الأجنبية كافة فروع المعرفة إلا أنها تستخدم بشكل أكثر في العلوم البحثية والتطبيقية.

ومن المؤكد أن النمو المعرفي وازدياد تشعب الحقول يفضيان إلى دخول مزيد من المفاهيم الجديدة إلى اللغة العربية سواء استطاعت اللغة استيعاب المفهوم الحديث بألفاظ عربية أم لا . والمتابع للدوريات العربية المتخصصة والكتب المترجمة يلحظ على الدوام دخول مصطلحات أجنبية دون أن يكون لها مقابلات عربية ثابتة . ولا شك أن بعض اللغات الأجنبية كالإنجليزية مثلاً أدق وأشمل من حيث الدلالة وضبط المصطلحات في الحقول العلمية سواء اللغة الطبيعية أو اللغات الاصطناعية التي تدعم وفرة الإنتاج الفكري وتوالد المصطلحات في مختلف الحقول .

ولو أخذنا بعض المفاهيم العلمية مثل (الليزر) وهو مفهوم مفرد لوجدنا أن المؤلفات عنه بالعربية لا تتجاوز العشرات من الكتب والمقالات الأساسية التي يتركز أغلبها على التطبيقات العامة لأشعة الليزر يقابلها أقل من عشرة مصطلحات أو مفاهيم مركبة باللغة العربية . بينما نجد في اللغة الإنجليزية آلاف الوثائق المتنوعة حول الموضوع يقابلها مئات المصطلحات المركبة والمقننة لتغطية مختلف جوانب مفهوم (الليزر) وتفريعاته وتطبيقاته المتنوعة التي تكاد أن تدخل في كافة حقول المعرفة المستقلة مثل الطب وما يتفرع عنه من مفاهيم كالعلاج والجراحة والعيون والأسنان والأنسجة والأوعية الدموية وكذلك موضوعات الصناعة والاتصالات والطيران والبصريات والمعلومات ومعالجة البيانات والطباعة والكيمياء والفيزياء والهندسة والعلوم الحربية والتصميم والفنون الجميلة وما يتجزأ من هذه الموضوعات العامة . هذا إلى جانب التفريعات الوجهية والشكلية للمفاهيم المنبثقة عن أشعة الليزر مثل :

الضوضاء ، النبض ، البقع ، انتشار الأشعة ، الأجهزة ، القياس ، التجارب والتحليل ، المواصفات ، القوانين ، المؤثرات ونحو ذلك من مقيدات وجهية وشكلية وجغرافية . والواقع أن الذي ينظر إلى مفهوم (الليزر) يحسبه مفهوماً دقيقاً ومخصصاً باللغة العربية ، وهذا يمكن ملاحظته بالقياس على كمية الإنتاج الفكري وعدد رؤوس الموضوعات المخصصة له بالعربية . بينما موضوع (الليزر) حقل واسع في اللغة الإنجليزية لدرجة أن المفهوم المفرد يتضاعف بالتركيب مع مفاهيم أخرى كثيرة وفي مجالات علمية متنوعة مع احتمال أن تكون مجالات تطبيق المفهوم متفاوتة في التدرج الهرمي صاعدة أو نازلة ومتساوية أو متقاربة داخل حقول علمية مختلفة .

وبمقارنة اللغة الطبيعية باللغات الاصطناعية نجد أن المكانز وقوائم رؤوس الموضوعات العربية لا تخلو من المصطلحات الأجنبية سواء تلك التي استقرت في الاستخدام منذ أمد أو تلك المصطلحات التي يحال منها إلى مقابلات عربية سواء كان المصطلح العربي شائع الاستخدام أم كان المصطلح باللغة الأجنبية ، بل إن استخدام المصطلحات الأجنبية في اللغات الاصطناعية يزداد كلما تعمقت درجة التخصيص واتسعت التغطية الموضوعية كما نرى ذلك بمقارنة (قائمة رؤوس الموضوعات العربية الكبرى) بغيرها من قوائم رؤوس الموضوعات العربية الأخرى .

ورغم أن المصطلحات باللغات الاصطناعية قد تكون أثبت وأكثر إحكامًا إلا أن المصطلحات الأجنبية قد تكون أدق في الاسترجاع وأكثر دلالة على المفهوم المطلوب من وجهة نظر الباحث في اللغة الطبيعية . والمثال على ذلك استخدام أسماء لغات وبرامج الحاسوب وكذلك العناصر والمركبات الكيميائية والمصطلحات الطبية الدقيقة ومسميات الأجهزة والآلات التي لها دلالات مفهومية مثل أجهزة القياس والمختبرات ، بل إن بعض المصطلحات الاجنبية تستخدم أكثر في المقابلات العربية إن وجدت . غير أن صعوبة التحكم بلغة الاسترجاع تكمن في طغيان المصطلح الأجنبي على العربي أو تساوي الطرفين في الاستخدام باللغة الطبيعية . كما أن استمرار دخول مصطلحات أجنبية عديدة يجعل اللغة الطبيعية أنسب للبحث عن المفاهيم الحديثة واسترجاعها قبل أن يثبت لها مقابلات عربية باللغة الاصطناعية المستخدمة في نظام الاسترجاع.

ومن المشكلات البارزة في استخدام المصطلحات الأجنبية المنطوقة بالحروف العربية اختلاف إملاء وتهجئة الكلمات واختلاف تركيب المفاهيم وتصريفاتها

ودمجها مع كلمات عربية أصلية ، هذا إلى جانب اختلاف درجة انتشار المصطلحات في بيئات علمية وعربية مختلفة . كما يتضح من تفاوت رسم المصطلحات الأجنبية اختلاف النطق بين البلاد العربية وفي استخدام حروف العلة أو تركها واختلاف النقحرة في بعض الحروف مثل (ج ، غ) (س ، ز) (ت، ط) . ونحو ذلك .

ونورد فيما يلي قائمة لبعض المصطلحات الأجنبية المستمدة من اللغة الطبيعية والاصطناعية والتي تغطي موضوعات متنوعة ومفاهيم مفردة أو مركبة ومتفاوتة في مستوى الخصوصية :

إلكترون ، إلكترونات ، إلكترونية
باسيك ، بيسك ، البازيك ، البيزك
دي بيس ، ديبيس ، دي بيز
الباسكال
الفورتران
كوبول
برنامج وورد برفكت
وورد ستار
وندوز
برنامج هارفارد
برنامج إكسل
برنامج الأوتوكاد
لغة لوجو ، لوغو
الكمبيوتر ، الكومبيوتر ، الكمبيترية ، الكمبيوترات

لوتس

التلسكوب

الميكروسكوب

الرادار

البولتكنيكي

الأكلينيكي

السيبرنتيك ، السيبر نطيقا

السيكوسوماتية

السوسيولوجيا ، سيسبولوجيا

الاتنولوجيا

باتولوجيا

أيديولوجيا

الديالتيكية

الأيدمولجيات

الفيبر جلاس

الأنيموجرافيا

الأوبريت

الأويتوفون

الأوينومتر

الأوسيلوجراف

الأينيسكوب

الأكتينومتر

الأكروماتية

أكسيد الديتوتروم

الأكورديون

التراميكروسكوب

البتربروتينات

البروبابلين

البريليوم

التيتانيوم

الثوريوم

الجشطلت

الجيروسكوب

الساماديوم

أشعة جاما

أشعة إكس

الإشعاع السنكروتوني

الملاريا

كاميرات الفديو

الجيتار ، الغيتار

الجودو

الكراتيه

البيروقراطية ، البيروقراطي

الدكتاتورية ، الديكتاتورية

الانتروبولوجيا ، الانثروبولوجيا

أوتاماتية ، أوتوماتيكية

الابستمولوجيا

البارامترية

البورجوازية

الاستراتيجية ، الاستراتيجيات

الأيكولوجيا

الأوكسينات

الجيوكيمياء

الديورايت

التوناليت

الثايرستور

الترياك

الديناميكا الكيميائية

الكيمياء

الأورماتية

الفلوروكوربون

الجيولوجيا الفيزيائية

جيومورفولوجية

البيدولوجي

الاوزون

هيدرولوجية ، الهيدرولية

فسيولوجيا

بروتين

الميكروبيولوجيا

نيماتولوجيا

الفيروس ، الفيروسات

الايدز

الهرمون

السوسيوانثربولجي

سيكولوجي

الجيوتكنيكية

الماكينات

الكهروميكانيكيات ، الكهروميكانيكية

الميتالورجيا

الفوتوجرافي ، الفوتوغرافي

الميكروموجية

الكهرومغناطيس

تكنولوجيا البيتوتون

الكلولسترول

الروماتيزم

الأنيميا

الماكياج

الكيك ، الكاتوه ، التورتات

التريكو

الكروشية

الأدراج البيونية

البلاتين

هيدروجين

كربون

التبولوجيا

الميكانيكا النيوتينية

ديناميكا

ستاتيكا

الفوتونات

الطبوغرافيا

البيوتكنولوجيات

الديموغرافيا ، الديموجرافيا

الباراسيكولوجي

الميتافيزيقيا ، الميكافيزيكا

الاستجولوجيا

الانطولوجيا

الميكروبروسور

ورغم أن هذه القائمة بمثابة عينة لأبرز المصطلحات الأجنبية المستخدمة في لغة المعلومات العربية إلا أنها تظهر نوعية المصطلحات ومقدار التفاوت في دلالة وتركيب المفاهيم وإملاء المصطلحات سواء كانت المفاهيم شائعة أو غير

مألوفة إلا في البيئة الأكاديمية أو المتخصصة . وهذا الحصر لا يشمل آلاف المفاهيم الدقيقة التي تقع في ثنايا النصوص العلمية المنشورة في البحوث والمقالات والكتب المترجمة مما يحتوي على مصطلحات أجنبية ليس لها مقابلات عربية موحدة .

ولا شك أن مشكلة التعريب والمصطلحات تزيد من صعوبة التحكم بلغة استرجاع المعلومات باللغة العربية سواء كانت طبيعية أو اصطناعية .

اللغات الطبيعية والاصطناعية :

يجب أن نشير إلى أن أعظم الصعوبات التي تواجه جميع نظم استرجاع المعلومات هي عملية التحكم بلغة التحليل والاسترجاع ، بل إن لغة المعلومات محور اهتمام دراسات علم المعلومات . وقد أدى ذلك إلى كثير من البحوث والتجارب المستمرة في حقول المعلومات والذكاء الاصطناعي والنظم الخبيرة واللغات والرياضيات في محاولة للوصول إلى تصميم نظم أو لغات مقننة لتطابق الاستفسارات التي يطرحها المستفيدون في المكتبات أو في غيرها من نظم الاسترجاع .

واللغة المستخدمة في هذا الصدد هي لغة التحليل والاسترجاع المستخدمة في وصف المفاهيم والموضوعات في التكشيف والفهرسة الموضوعية . واللغة كما يعرفها Cleveland (هي الكلمات المتاحة لوصف وثيقة معينة أو المستخدمة لصياغة استفسار عند البحث عن وثيقة ، وذلك بما يشمل القواعد التي تصف كيفية استخدام المصطلحات والعلاقات فيما بينها).

وتنقسم لغات التكشيف أو لغات المعلومات إلى قسمين رئيسين هما :

اللغـة الاصطناعـيـة المحكمـة التي يتم تعـيـينهـا لمحـتـويات الوثائق خلال التحليل والاسترجاع . وقد تكون اللغة الاصطناعية رمزية مثل نظم التصنيف الرقمية والحرفية أو تكون لفظية مثل المصطلحات والمواصفات المستخدمة في المكانز وقوائم رؤوس الموضوعات . والنوع الآخر هي اللغة الطبيعية أو الحرة . وهي المفردات التي تشتق من الوثائق عند التكشيف والاسترجاع كما في كشافات التباديل أو كشافات الكلمات الدالة في السياق وكما في المستخلصات المكتوبة بلغة طبيعية .

ومع البدايات الأولى لاستخدام كشافات التباديل ظهر خلاف حاد بين أنصار التكشيف الموضوعي الذي يعتمد على واصفات موضوعية مقننة ، وبين دعاة استخدام مفردات اللغة الطبيعية المستمدة من المداخل أو النصوص المكشفة . ونتيجة لهذا الخلاف ظهرت مدرستان أو نظامان للتحليل والاسترجاع، أحدهما نظام الاسترجاع بالاعتماد على اللغات الاصطناعية والآخر نظام اللغة الطبيعية.

ولقد كان النظام الأول وما زال قويًا نظرًا لارتكازه على معايير ثابتة في التحليل والاسترجاع . وهذا ما يمكن أن نلمسه بوضوح في استمرارية تطوير واستخدام المكانز وقوائم رؤوس الموضوعات المتعددة . أما اللغة الطبيعية فتمتاز بالثراء والمرونة التي تزيد من صعوبة البحث والاسترجاع إلى جانب مميزات حيوية متعددة كما سنرى فيما بعد .

ويمكن تلخيص مميزات وعيوب اللغات الاصطناعية المحكمة نقلاً عن Aluri فيما يلي :

مميزات استخدام اللغات الاصطناعية :

١ - تقوي الاحتمال في أن يعبر كل من المكشف والباحث عن المفهوم بنفس الأسلوب مما يعزز عمليات المضاهاة .

٢ - تزيد من احتمال أن يستخدم المكشف وغيره من المكشفين المصطلح نفسه في أوقات مختلفة مما يضمن ثبات الاستخدام للمصطلحات .

٣ - يزيد استخدام اللغة المحكمة في احتمال توجيه كل من المكشف والباحث إلى موضوعات مرغوبة من خلال الإحالات إلى موضوعات متقاربة .

٤ - يساعد استخدام اللغة المحكمة في الإسراع بعملية البحث من حيث أن اللغة المحكمة تمكن الباحث من الجزم بأن ما يبحث عنه من مواد أو معلومات غير متاح وبأنه لا داعي للبحث تحت مترادفات متنوعة .

عيوب اللغات الاصطناعية :

١- إن اللغات الاصطناعية مهما بلغت قوة تركيبها فهي تظل متقادمة للاستخدام في التكشيف والاسترجاع .

٢ - تكفي النصوص العلمية باللغة الطبيعية للتكشيف والاسترجاع بشكل كامل .

٣ - إن التكشيف الآلي باللغات الطبيعية يواكب التطورات العلمية .

٤- يمكن صياغة الاستفسار في جملة أو عبارة والاسترجاع باللغة الطبيعية أثناء البحث الآلي المباشر

ورغم قوة حجج المؤيدين لاستخدام اللغات الاصطناعية إلا أن بعض الدراسات أظهرت أنه لا توجد فروق بارزة في نتائج الاسترجاع من حيث التحكم والاستدعاء . ويقول Cleveland في كتابه الذي يعتمد أساسًا على التحليل الموضوعي واستخدام اللغات الاصطناعية : " ... بأن المؤيدين لنظم اللغات المحكمة يرون بأنها أكثر ثباتاً وأكثر فاعلية في الاسترجاع ، إلا أن التجارب التي تمت خلال السنوات الماضية قد فشلت في الوصول إلى نتائج مقنعة " أي مقنعة بوجهة نظر المؤيدين لنظم اللغات الاصطناعية . أما Lancaster فيرى بأن أبرز التوجهات المهمة في السنوات الأخيرة هي في الميل نحو استخدام اللغات الطبيعية في التكشيف بدلاً من اللغات المحكمة . كما أظهرت البحوث خلال عشرين سنة الماضية أن هناك تحولاً في التفكير إلى أن أساليب الاسترجاع المتعددة تتكامل مع بعضها البعض وليست متضاربة أو متضادة كما يقول Travis و Fidle في ختام دراستهما المسحية عن أدبيات التكشيف ولغاته المستخدمة في التحليل الموضوعي .

خصائص الاسترجاع باللغة الطبيعية :

يجب أن نشير إلى أن اللغة الطبيعية المقصودة هي لغة المعلومات المستخدمة في التعبير عن المفاهيم الموضوعية عبر المداخل من عناوين وكشافات ومستخلصات . وتقوم كشافات الكلمات الدالة بوظيفة الاسترجاع عبر الواصفات الموضوعية الموجودة باللغة الطبيعية التي استخدمها المؤلفون في العناوين أو في المستخلصات الحرة ؛ بل إن إعداد كشافات التباديل يتم تلقائيا بعد تكوين السجل الببليوجرافي لأي وثيقة . وعدا السرعة في إعداد الكشافات فإن محاسن استخدامها يعتمد بالدرجة الأولى على قدرات نظام

الاسترجاع في فرز وترتيب الكلمات المفتاحية التي قد تكون ذات دلالات مهمة في الاسترجاع الموفق والمطابق لخصوصية البحث بدرجة تفوق اللغات الاصطناعية من مكانز ورؤوس موضوعات . ومعنى ذلك أن تصميم خطة البحث بالاعتماد على الكلمات المستخدمة في العناوين وبالذات عناوين البحوث والمقالات المتخصصة أكثر دقة وأقرب لاحتياجات الباحثين من الاعتماد على رؤوس الموضوعات العربية التي هي في الغالب قوائم عامة أو شبه عامة .

وفيما يختص باللغة العربية نجد أن اللغات الاصطناعية المتاحة متقادمة، ولا تتضمن في الغالب الكثير من المفاهيم الضيقة والتى يمكن أن تظهر في الإنتاج الفكري المتخصص والحديث ، هذا بالاضافة إلى أن المكانز ورؤوس الموضوعات لا تتجه نحو حصر الأسماء والأشياء التي تظهر من خلال عناوين البحوث والتقارير ومقالات الدوريات الجارية .

والواقع أن لغة العناوين أقرب في نقل محتويات الوثائق الأصلية في أغلب الموضوعات العلمية والتقنية والحقول المتخصصة . والسبب في ذلك يعود إلى أنه لا يوجد لغة أو محلل معلومات وسيط بين الباحث ولغة الوثيقة الأصلية فاللغة الوسيطة قد تكون قاصرة ، كما أن المكشف أو محلل المعلومات قد يقع في أخطاء كثيرة أثناء التحليل الموضوعي . فقد يكون عنوان المقال مثلا «صناعة حامض الاستيك في السعودية » بينما قائمة رؤوس الموضوعات المستخدمة تتضمن مصطلحاً واسعاً هو (الصناعة البتروكيمياوية) . أو قد يعمد المكشف إلى الخطأ أو الاجتهاد واستخدام المصطلح الأوسع بدلا من الموضوع الدقيق والمخصص الذي سيظهر في كشاف التباديل دون تدخل من

المكشف . ويمكن قياس مستوى الخصوصية ونسب التحقيق والاستدعاء باللغة الاصطناعية من خلال المقارنة بين عناوين المقالات والبحوث مع رؤوس الموضوعات المعطاة لها في نظام الاسترجاع (الصوينع) .

كما تتضح عيوب اللغات الاصطناعية في المكتبات التي تعامل فيها كافة أوعية المعلومات على حد سواء في التحليل الموضوعي . ولذا تتساوى موضوعات الكتب والبحوث والمقالات والتقارير في مستويات الخصوصية مما يؤدي إلى زيادة نسبة الاستدعاء على حساب التحقيق . بينما تقوم اللغة الطبيعية بالإفصاح عن مستويات الخصوصية الموجودة في لغة العناوين أو المستخلصات بشكل أدق للموضوعات المتخصصة والأضيق .

وعلى الرغم من العيوب والمشكلات الدلالية والتركيبية للغة الطبيعية إلا أن استخدامها للبحث والاسترجاع يظل أفضل من استخدام المكانز ورؤوس الموضوعات العربية وبالذات للاحتياجات المعلوماتية الدقيقة في المجالات الموضوعية المتجددة بسرعة سواء في العلوم الاجتماعية أو التطبيقية . والسبب في ذلك أن اللغات الاصطناعية متخلفة عن حركة التأليف والترجمة باللغة العربية، كما أن تحديث المكانز ورؤوس الموضوعات لا يجاري حركة الإنتاج الفكري المتنامي . ولا يقتصر الأمر على موضوعات فرعية من جزئيات المعرفة، بل إن هناك حقولا كاملة ينقصها التغطية المصطلحية الكافية والمقننة ، وهذا ما يجعل هناك فجوة زمنية تؤدي خلالها اللغة الطبيعية دوراً مهمًا في تلبية الاحتياجات البحثية للمستفيدين الذين ينشدون المعلومات الأحدث .

واللغات الاصطناعية بحد ذاتها متفاوتة في توحيد المصطلحات على المستوى الوطني وعلى المستوى العربي سواء في تقنين المفاهيم المفردة أو المركبة

وصياغتها وهذا الأمر يرجع إلى المشكلة المزمنة في التعريب وتوحيد المصطلحات باللغة العربية . ورغم أن الرغبة في توحيد المصطلحات قد تكون مدعاة لتفضيل اللغات المحكمة إلا أن تحقيق ذلك عمليًا قد لا يجعل اللغة الاصطناعية تتفوق على اللغة الطبيعية في دقة الاسترجاع لعدة أسباب منها: اختلاف التخصصات والاهتمامات الموضوعية بين نظم المعلومات وعدم وجود مكانز كاملة وحديثة أو مواكبة للتطورات العلمية في مختلف حقول المعرفة . ورغم أن بعض المؤسسات المعلوماتية العربية تقوم بتطوير المكانز والمصطلحات الخاصة بها إلا أنه يصعب في الوقت الراهن تحقيق التوافق بين نظم الاسترجاع المختلفة باستخدام اللغات الاصطناعية المتوفرة باللغة العربية .

والحقيقة أن خلفيات الباحثين وسلوكيات البحث والثقافة المحلية وحركة التأليف والترجمة في البيئات العربية المختلفة تجعل الاسترجاع باللغة الطبيعية أقرب إلى سلوكيات الباحثين خصوصا إذا علمنا أن المفاهيم تولد باللغة الطبيعية أولاً ثم يتداولها الباحثون في الاتصال المعلوماتي قبل تثبيتها في المعاجم واللغات الاصطناعية بوقت طويل . ويقول (عبدالهادي) : " يجب استخدام رأس الموضوع الشائع ، فالاستخدام في المكتبة المصرية يعني الاستخدام المصري الجاري ، وإذا لم يطبق هذا المبدأ فإن معظم القراء سوف لا يجدون المواد التي يرغبونها تحت الرأس الذي يخطر على بالهم أولاً " .

ولقد اعتمدت اللغات الاصطناعية في صياغة المفاهيم باللغة العربية على الأدبيات الأجنبية والقواعد المطبقة في قائمة رؤوس موضوعات مكتبة الكونجرس أو قائمة سيزر (العناني) .

ومن القوائم المعيارية المستخدمة في المكتبات العربية :

قائمة رؤوس الموضوعات العربية الكبرى ١٩٨٥م. شعبان خليفة ـ الرياض .

قائمة رؤوس الموضوعات العربية ١٩٩٣م . معهد الإدارة العامة ـ الرياض .

قائمة رؤوس الموضوعات العربية ١٩٧٨م . إبراهيم الخازندار - الكويت .

رؤوس الموضوعات العربية ١٩٧٨م - جامعة الرياض .

قائمة رؤوس الموضوعات المالية والاقتصادية ١٩٩٠م . المركز الوطني للمعلومات المالية ـ الرياض .

ورغم أن اللغات الاصطناعية العربية ما زالت تعاني بعض المشكلات في التعريب والصياغة والشمول والحداثة إلا أنها قد حظيت بعناية ودراسات متكاملة بخلاف اللغة الطبيعية ومشكلاتها ومقارنتها باللغات المحكمة خصوصا بعد أن أصبحت اللغة الطبيعية متاحة للاسترجاع بشكل أوسع .

والخلاصة : أن كفاءة الاسترجاع تعتمد في النهاية على قدرات ومهارات الباحث أو متخصص المعلومات في السيطرة على لغة المعلومات وتغطية المفاهيم المستهدفة بالمصطلحات والتراكيب المتغيرة في اللغة الطبيعية أو الاصطناعية عند إعداد خطة البحث .

ومهما قيل في المفاضلة بين اللغة الطبيعية والاصطناعية فإن لكل منهما محاسنه وعيوبه إلا أنهما معا يعضدان عملية الاسترجاع ويتيحان إمكانات واسعة ومتعددة للتعامل مع المعلومات حسب التطورات في نظم المكتبات والمعلومات .

مشكلات الاسترجاع باللغة الطبيعية :

على الرغم من سهولة استخدام اللغة الطبيعية في المستخلصات وفي كشافات الكلمات الدالة في سياق العناوين ، إلا أن هناك صعوبات كثيرة ، بعضها تقليدية وعامة أي ترجع إلى مرونة اللغات الطبيعية وتعدد وتفاوت مفرداتها في التعبير عن المفاهيم ، وبعضها مشكلات صرفية أو نحوية تختص باللغة العربية . فالعربية تمتاز بحدة الخاصية الصرفية وحرية ترتيب الكلمات داخل الجمل إلى جانب اللبس الناجم عن غياب التشكيل ونحو ذلك كما يقول (نبيل علي) . ولا شك أن الاعتماد الكلي لإنتاج الكشافات على اللغة الطبيعية بمعالجة الحاسوب قد يكون له جوانب سلبية . ويعيق فاعلية الاعتماد على اللغة الطبيعية في نظام المعلومات مشكلات دلالية وتركيبية واسترجاعية يمكن إجمالها كما يلي :

١ - التصريف والاشتقاق :

تمتاز اللغة العربية بالثراء في التصريف والاشتقاق ، إلا أن التغيرات الصرفية المعقدة التي تطرأ على الكلمات تقع في كثير من الأحيان في أوائل الكلمات مما يشتت المداخل المتشابهة ويعيق البحث والاسترجاع باستخدام روابط الصلة المنطقية (و ، أو ، غير) أو إعداد خطة بحث محكمة بالاعتماد على الكلمات المبتورة . فهذه الأنماط من البحث تعتمد على استخدام مصطلحات محددة يمكن ربطها والجمع بين مفاهيم متعددة حسب خطة البحث والربط اللاحق للمفاهيم المفردة .

فإذا كان مصطلح البحث كلمة (إدارة) مثلا فإن لهذه الكلمة تصاريف وصيغاً إعرابية متعددة يؤثر تفاوتها على البحث والاسترجاع مثل التصاريف المغايرة التالية :

إدارات ، إداري ، إدارية ، إداريون ، إدرايين ، إداريات ، مـدير ، مـديرة ، مدراء ، مديرات ، مديرية ، مديرو ، مديري ، يدير ، تدير .

وبعض هذه التصاريف قد يلحقها زوائد متعددة مثل حروف الجر المتصلة وأل التعريف (بالإدارة ، للإدارة ، كالإدارة) ونحو ذلك من تصاريف أخرى . وهذا يؤدي إلى زيادة احتمالات تشتت المفهوم الواحد في مواقع ألفبائية متعددة من كشافات التباديل ، ومن ثم سيتم إغفال الكلمات الأخرى التي لا تتطابق تماما مع مصطلح البحث الذي اختاره الباحث خلال الاسترجاع . وعلى الرغم من إمكانية القيام بخطط بحث تجمع مختلف تصاريف الكلمات الدالة إلا أن ذلك لا يعد مرونة في نظام الاسترجاع لأنه يستهلك وقتاً طويلا للتفكير في جميع الاحتمالات المتوقعة للصياغات المغايرة لمصطلح البحث المجرد من الزوائد الذي تم استخدامه في بداية البحث . كما أن ذلك يلقي مزيدا من أعباء البحث على المستفيد الذي قد لا تكون لديه خبرة كافية لا في مصطلحات الموضوع ولا في مهارات البحث وأساليبه .

ولعل مشكلة تشتت المداخل واختلاف ترتيب الكلمات الدالة على مفهوم واحد تتضح في التفاوت بين شكل الكلمات في صيغتي المفرد والجمع مما يستوجب البحث تحت مصطلحين دالين على مفهوم واحد مثل :

طفل ، أطفال

مال ، أموال

سلاح ، أسلحة

سمك ، أسماك

حمض ، أحماض ، حمضيات ، الحوامض .

٢ - المترادفات :

تكثر في اللغة العربية الكلمات المترادفة الدالة على مفهوم واحد . فالمعنى الواحد قد يأخذ أشكالاً متعددة من الكلمات . ومشكلة الترادف في الاسترجاع أنه يؤدي إلى توزيع الكلمات التي تحمل المفهوم نفسه في مواقع متعددة من الكشافات . وهذا أشد خطورة من تشتيت الكلمات المتفرقة من جذر واحد كما رأينا سابقاً . فالمداخل التي يفترض فيها أن تكون تحت رأس واحد تتشتت في أماكن متعددة قد لا يتوقعها الباحث . فالباحث الذي يريد استرجاع كلمة (قوانين) قد لا يعرف بأن هناك كلمات أخرى للمفهوم نفسه استخدمت في العناوين أو المستخلصات مثل (تشريعات ، أنظمة ، أحكام) ونحوها . وفي اللغة العربية الكثير من المترادفات التي ترجع إلى عوامل كثيرة مثل :

حقول البترول ، آبار البترول ، أو حقول وآبار النفط .

الأسرة ، العائلة .

بنوك ، مصارف .

الحاسوب ، الكمبيوتر .

والحقيقة أن مشكلة المترادفات محلولة في المكانز ولكنها في اللغة الطبيعية كثيرة ومتفاوتة حسب الحقول العلمية واختلاف مصادر النشر والتعريب مما يعكس مشكلة المصطلح واختلاف البيئات العلمية في البلاد العربية .

٣ - الألفاظ المشتركة :

الألفاظ المشتركة أو المتجانسة حيث تأخذ الكلمة الواحدة في اللغة الطبيعية عدة معان ومفاهيم بحيث لا يمكن تمييز المعنى الصحيح أو المستهدف

للكلمة إلا في السياق . والمشترك اللفظي كما يقول (حلمي خليل) :

" هو اللفظ الواحد الدال على معنيين مختلفين فأكثر " . ورغم أن إشكالية المشترك اللفظي في لغة المعلومات ليست حادة كما في نظم المعلومات النصوصية إلا أنه قد يرد في المستخلصات وفي عناوين الكتب العربية كلمات كثيرة مفردة أو مركبة لها دلالات مختلفة مثل :

الملكية (نظام الحكم)

الملكية (قانون)

الدين (مثل الإسلام)

الدين (القرض)

العلم (الحقل والتخصص)

العلم (الراية)

سنة (مذهب)

سنة (عام أو حول)

الفاشية (المنتشرة)

الفاشية (في السياسة)

فهذه الكلمات من المشترك اللفظي على اعتبار أنه لا يوجد في نظام الاسترجاع علامات للتشكيل والحركات بحيث يمكن التمييز بين الكلمات المشتركة في اللفظ والمختلفة في المعنى بحسب نطقها خصوصا في حالات البحث والاسترجاع بالمفاهيم المفردة . ويبدو أن تأثير السياق في فهم دلالة الكلمة العربية قوي لدرجة أنه قد لا يمكن الاعتماد على المفاهيم المفردة في الاسترجاع أحياناً ، وإنما لا بد أن يعرف الباحث موضع الكلمة ضمن سياقها إذا

لم تكن المفاهيم مركبة . وهذا ما يدعو إلى ضرورة إضافة مقيدات بعد الألفاظ المتجانسة لتفسير دلالتها كما يقول (فوسكت) .

٤ - التركيب :

بما أن المفاهيم المركبة أكثر استخدامًا في اللغات الاصطناعية والطبيعية وأكثر تنوعاً في الصياغات اللغوية فإننا نجد أن الكلمات المتلازمة تمثل بعض الصعوبات في البحث والاسترجاع باللغة الطبيعية . ويمكن القول إجمالاً إن المشكلات اللغوية التي تعتور المفهوم المفرد من حيث الصياغة قد تنتقل إلى المفاهيم المركبة غير أن التلازم اللفظي بين المفاهيم يلغي أو يضعف اللبس الدلالي ، إما بالوصف أو التخصيص كما في الربط المسبق باللغة الاصطناعية في حالات المضاف والمضاف إليه والصفة والموصوف والمفاهيم المربوطة بحروف الجر .

ويغض النظر عن مشكلات التركيب الاصطناعي ، فإن الاسترجاع بالمفاهيم المركبة يعتمد على دمج المفاهيم المفردة بالربط اللاحق . وهذا ما يخفف كثيرا من المشكلات التقليدية مثل قلب الرأس والإحالات بين الكلمة الأولى والثانية، أو استخدام الروابط النحوية غير أن مرونة التركيب الطبيعي قد تثير بعض المشكلات في الاسترجاع . ومن ناحية التركيب فإن أنسب الصيغ للبحث والاسترجاع هي المفاهيم المركبة التي يقل تأثير التركيب اللغوي عليها سواء من ناحية الصياغة أو الإعراب وذلك يشمل المفاهيم المفردة المربوطة بحروف الجر المنفصلة والمضاف والمضاف إليه المكون من كلمات دالة مفردة مثل :

الطب عند العرب

الفلسفة في الإسلام

التأمين ضد الحريق

تعليم المرأة

صحة الطفل

ومع ذلك لا يمكن التحكم في دلالات المفاهيم المركبة في اللغة الطبيعية . فالمضاف والمضاف إليه قد يتبادلان الموقع الإعرابي فتتغير دلالة المفهوم المركب بسبب الارتباط المزيف بين المفاهيم مما يؤدي إلى ضعف التحقيق وزيادة نسبة الاستدعاء عند الاسترجاع بالربط اللاحق للمفاهيم المستقلة مثل :

إدارة العلوم

علوم الإدارة

بحوث المكتبات

مكتبات البحوث

تصنيف الوظائف في المكتبات

وظائف التصنيف في المكتبات

ميزانية الوظائف الحكومية

وظائف الميزانية في الحكومة

لذا نجد أن دلالة المفاهيم المترابطة تتغير حسب التركيب وترتيب المفاهيم أو الموقع الإعرابي سواء كانت الكلمة الدالة مضافًا أو مضافًا إليه، إلا إذا كان لترتيب الجملة البحثية وموقع الرابط المنطقي بين المفاهيم المتجاورة تأثير في فرز الكلمات المتلازمة حسب موضعها في السياق أو النص كما في بعض نظم استرجاع المعلومات .

وإذا كـان الاسـترجاع يتـأثر بالحـالة الإعـرابيـة المتـغيـرة فإن ذلك يستـدعي التفكير في خطة البحث . وعند الاسترجاع لا يمكن النظر إلى المفاهيم على أنها مسـتـقلة دائـما . فـإذا كـان المفـهوم الموصـوف اسـماً مـفـردا فـهـو يحـتـاج إلى التخصيص النعتي المفرد أو ما يوافق الجمع إذا كان جمعًا مثل :

القانون الإداري

القوانين الإدارية

مع احتمال أن يأتي المفهوم المركب ذاته في صياغة أخرى وفي سياق مختلف على أنه مضـاف ومضـاف إليه مثل (قانون الإدارة) أو في تركيب طبيعي آخر مثل (القـانون وأثر تطبيـقه على الإدارة) وذلك يستلزم وضع خطة البحث على هذا النمط .

(القانون أو القوانين والإداري والإدارة والإدارية)

هذا عـلاوة على تفـاوت المفـاهيم المترادفـة باخـتـلاف التـركـيب أو الصـيغ واستـخدام لغة مرنة للتـعبير عن المفهوم المركب دون النظر للمصطلحات السائدة في الإنتاج الفكري المتخصص مثل :

اقتصاديات الدول الآخذة في النمو = الدول النامية

التحليل الإحصائي = الإحصاء التحليلي

التوعية الإسلامية = الوعظ والإرشاد = الأمر بالمعروف والنهي عن المنكر.

٥ - صيغة الفعل :

المفاهيم لا تأتي إلا عـلى هيئة أسماء أو صفـات في اللغات الاصطناعيـة أما في اللغـة الطبيـعـيـة ، فإن المفاهيم قد تأتي على صيغ أفـعال في بعض الكتب التعليمية والإرشادية وكتب الأطفال مثل العناوين التالية :

كيف ندعو الناس ؟

كيف تحج أيها المسلم ؟

كيف تزكي أموالك ؟

كيف تصلي ؟

كيف تصوم ؟

هكذا نصوم

فالكلمات الدالة في العناوين هي المفاهيم المشتقة من الأفعال (ندعو = الدعوة ، تحج = الحج ، تزكي = الزكاة ، تصلي =الصلاة ، نصوم = الصيام). فالباحث الذى يرغب في كتب عن المفاهيم المذكورة قد لا يتوقع بأنها تجيء في صيغ الأفعال باللغة الطبيعية . كما تأتي كتب الأطفال في صيغ فعل أمر مثل (اقرأ واكتب ، اكتب ولوّن) . لتدل على مفاهيم (القراءة ، الكتابة ، التلوين أو الرسم) .

٦- الإملاء ورسم الكلمات :

بغض النظر عن الحساسية الدلالية والمعجمية للألفاظ بسبب غياب أو نقص حركات التشكيل من الفتحة والضمة والكسرة والشدة وكذلك تفاوت إملاء الكلمات المبدوءة بهمزة في النصوص العربية ، فإن لمعضلة تفاوت رسم الكلمات جوانب سلبية على استرجاع المعلومات .

وترجع ظاهرة تفاوت الإملاء إلى سببين ، أحدهما الأخطاء البشرية ، وثانيهما اختلاف الممارسات أو القواعد المتبعة لإملاء الكلمات المعربة والأسماء الأجنبية والتي يختلف رسمها بين المؤلفين العرب ، مما يؤثر على دقة

الاسترجاع باللغة الطبيعية كما في الكلمات التالية :

ببليوجرافيا ، ببليوغرافيا

مكروفلم ، ميكروفلم

كمبيوتر ، كومبيوتر

انجلترا ، انكلترا

رومانتيكية ، رومانسية ، رومانطيقية

وسيتم التطرق لتأثير الإملاء على استرجاع المعلومات في مواضع متعددة من الكتاب .

٧ - الرمز والمجاز :

استرجاع المعلومات مرتبط بالمفاهيم الدالة والموضوعات إذا تجسدت ضمن ألفاظ اسمية مستقلة ودالة على كيانات أو عمليات أو مجردات في كافة حقول المعرفة ، فإذا كانت اللغة الطبيعية خالية من المصطلحات أو الكلمات الدالة فإن الاعتماد على العنوان وحده في البحث قد لا يوصل إلى المعلومات المطلوبة بسبب الانفصام بين نص العنوان والمحتوى الموضوعي للكتب ، ويتضح ذلك في عناوين كثيرة في اللغة العربية مثل :

معالم على الطريق	سيد قطب
الإطراف بأوهام الأطراف	ولي الدين العراقي
شرح قطر الندى وبل الصدى	ابن هشام الأنصاري
كشف الستار عما خفي من أسرار	عدنان تللو

وهناك من يقول بأن العنوان ليس مدخلا مثاليًا لأي كشاف ، لأن مؤلفي الكتب والوثائق لا يضعون في حسبانهم عند صياغة العناوين بأنها سوف تستخدم لأغراض التكشيف والاسترجاع . لذا فإن بعض العناوين قد لاتتضمن من الكلمات ما يكفي للدلالة على محتوى الوثيقة إما بسبب استخدام المجاز والرمز أو باستخدام مفاهيم لا علاقة لها بنص الوثيقة ، كما أن بعض المؤلفين يستخدمون عناوين غريبة أو مبهمة وخاصة في الموضوعات الأدبية والفنية أو كتب التراث القديمة والمسجوعة . ولا شك أن قيمة التباديل والاسترجاع تعتمد كثيرا على مستوى تمثيل كلمات العنوان لمضمون الوثيقة المكشفة ، وهذا ما يجعل تطبيق تكشيق الكلمات ينجح في الموضوعات العلمية والتقنية أو المجالات البحثية مثل النثر الفني والشعر والقصص ونحوها فإن كشافات الكلمات غير مفيدة . فمثلا عنوان ديوان (أنشودة المطر) لبدر شاكر السياب لا علاقة له بعلم المناخ والطقس ، كما أن قصص وكتب توفيق الحكيم (زهرة العمر، يا طالع الشجرة ، شجرة الحكم ، أشواك السلام) لا علاقة لها بالنبات والزراعة رغم وجود كلمات (زهرة ، أشواك ، شجر) .

٨ - التحقيق والاستدعاء :

إن مشكلات اللغة الطبيعية المتعلقة بدلالات المفاهيم وتركيبها تؤثر في دقة الاسترجاع وسرعة الوصول إلى المعلومات المستهدفة بالشكل المناسب من حيث الجودة والكمية ، ويتعلق التحقيق بنوعية المعلومات الصالحة للمفاهيم والمصطلحات المستخدمة في البحث ، بينما يعني الاستدعاء كمية المعلومات المسترجعة والتي قد تنقص أو تزيد عن الاحتياجات البحثية . فإذا تجاوزت

المعلومات المسترجعة حاجة المستفيد وزادت نسبة الاستدعاء ، فلأنها تتضمن معلومات حشو أو شوشرة معيقة للوصول الفعال إلى المعلومات المناسبة . ومن العوامل المؤثرة بالتحقيق والاستدعاء تفاوت التصاريف والمترادفات والمشترك اللفظي والتركيب اللغوي المتغير والإملاء . كما أن من العوامل التي تضعف التحقيق وترفع نسبة الاستدعاء المشكلات المنهجية في التأليف حيث يلجأ بعض المؤلفين إلى استخدام بعض الكلمات العامة التي تساير التوجهات الإعلامية السائدة في كافة مجالات الحياة العامة والسياسية والاقتصادية ، ومن الكلمات المستخدمة على نطاق واسع كلمة (التنمية) التي يتم ربطها مع مفاهيم متخصصة ، وعند الإمعان في المحتوى الموضوعي للمؤلفات نجد أن كلمة (التنمية) استخدمت لجذب القارىء أكثر من كونها معالجة جادة للموضوع المركب ، ومن أمثلة ذلك (المكتبات والتنمية ، الحاسب الآلي والتنمية) ونحو ذلك وعند تحليل الموضوع نجد أن المعالجة آحادية الجانب ، أي ليس في ثنايا الكتاب ما يعكس دلالة العناوين بدقة وإنما تم التركيز على الجانب الفني للموضوع دون معالجة لعلاقة التفاعل بين (التنمية) والموضوع الآخر ، والعناوين المضللة كثيرة إما عن قصد أو غير قصد ، ولكنها في النهاية تعيق الاسترجاع والوصول الفعال للمعلومات المطلوبة .

هذه أبرز المشكلات اللغوية الخاصة باستخدام اللغة الطبيعية التي يمكن ملاحظتها على بعض الكلمات والتراكيب في العناوين العربية . ولا شك أن هناك مشكلات دلالية واسترجاعية أخرى تتصل جميعها بتصميم نظام الاسترجاع والمعالجة اللغوية والمنطقية لكلمات التوقف وطرق عرض المعلومات .

الفصل الثالث

كلمات التوقف

- خصائص كلمات التوقف
- أنواع كلمات التوقف
  * الحروف والأدوات النحوية
  * المفردات غير الدالة
- تكوين كلمات التوقف



الفصل الثالث

كلمات التوقف

تعتمد قيمة الاسترجاع في الكشافات والمستخلصات على مستوى تمثيل كلمات العنوان أو المستخلص للمحتوى الموضوعي للوثيقة المكشفة . وبما أن اللغة الطبيعية تحتوي على كلمات كثيرة غير دالة ومتفاوتة الصياغات ، فإن نظام الاسترجاع يدعو إلى ضرورة تطويع اللغة الطبيعية في سبيل تسهيل الوصول إلى المعلومات في نظم المكتبات ، ولذا يجب أن نستبعد من سياقات العناوين أو المستخلصات الكثير من المفردات الفارغة من المعنى وغير الدلالية لا سيما إذا كانت تتردد في النصوص بصورة كبيرة ثم الإبقاء على الكلمات الدالة التي تحمل مضامين اصطلاحية وموضوعية أو مفاهيم مهمة فى الاسترجاع .

ويمكن تطويع اللغة الطبيعية وفق الأساليب الآلية واللغوية كما يلى :

١ - إنشاء قوائم توقف للمفردات غير الدالة في الاسترجاع .

٢ - وضع قواعد منطقية لمعالجة المفردات ذاتيا .

٣ - تحديد رموز خاصة على لوحة مفاتيح المطراف تستخدم في مرحلة الفهرسة الآلية وإدخال البيانات الببليوجرافية وتستخدم الرموز لاستبعاد أو استثناء المفردات حسب دلالتها في الحالات المختلفة .

أما قائمة التوقف أو قائمة الاستبعاد Stop list فهي مسرد للمفردات التي لا تحمل معاني اصطلاحية أو مفاهيم مهمة لأغراض استرجاع المعلومات ، ولذا يتم إهمالها آلياً من المستخلصات وكشافات التباديل بعد مضاهاة النصوص والعناوين على قائمة التوقف خلال الترتيب والاسترجاع. ويقابل قائمة التوقف ، قائمة أخرى تستخدم أحيانا للكلمات المهمة المبتغاة أو المجازة Go List ، وهي قائمة الكلمات المطلوب اختيارها دون سواها عند المضاهاة لتكون هي الكلمات المفتاحية المستخدمة في الاسترجاع . ويعتمد تحديد الكلمات المجازة في القدرة على السيطرة على الكلمات المهمة وذلك بتكوين مكنز أو قاموس للكلمات الدلالية المحتمل وقوعها في النصوص المختزنة ، ويلائم هذا الأسلوب نظم استرجاع النصوص المحدودة حيث يمكن السيطرة على لغة الاسترجاع .

أما البديل الآخر عن استخدام قائمة التوقف المحددة للمفردات غير الدالة فسيتم من خلال اعتماد أسلوب الكلمات المجازة أثناء التكشيف ، ويتم ذلك من خلال وضع الإشارة المعينة قبل الكلمة الدالة والمرغوبة في الاسترجاع كما في العناوين التالية :

تعليم اللغة العربية للأجانب

كيف تتعلم الأسبانية بدون معلم ؟

المشتقات في اللغة الانجليزية

محاسبة التكاليف واستخداماتها الإدارية

الإيدز الطاعون الجديد

الإجهاض بين الفقه والطب

من الذرة إلى الليزر

النظائر المشعة

مكتبة الملك فهد الوطنية

فالكلمات المرغوبة في التباديل والاسترجاع هي : (تعليم ، اللغة ، العربية، للأجانب ، الأسبانية ، المشتقات ، اللغة ، الإجليزية ، محاسبة ، التكاليف ، الإدرايه ، الإيدز ، الإجهاض ، الفقه ، الطب ، الذرة ، النظائر ، المشعة ، مكتبة ، الملك ، فهد ، الوطنية) . وما عدا الكلمات الدالة يستبعد إلى جانب الحروف الملتصقة بالكلمات الدالة مثل حروف الجر والعطف وأل التعريف على أن يتم ذلك من خلال التمرير التلقائي على قائمة التوقف الواضحة والمحدودة .

ورغم أن هذا الأسلوب أوثق من حيث تجنب المحاذير الدلالية لتمييز الكلمات حيث يمكن الحكم بسهولة على الكلمات الدالة داخل سياقاتها إلا أنه من حيث التطبيق يستغرق وقتًا طويلاً في الوقوف على كل كلمة دالة أثناء التكشيف على نماذج الإدخال من قبل المفهرس ومن قبل مدخل البيانات الذي عليه بعد إدخال النص المرور على الكلمات الدالة المستثناة ثم وضع الرموز المناسبة أمامها . كما أن استخدام هذا الأسلوب لا يعني بالضرورة التخلص من المشكلات الدلالية والتركيبية الملازمة للغة الطبيعية وبالذات العناوين .

ولذا فإن هذا الاسلوب أو بالتخصيص أسلوب وضع إشارة معينة على الكلمات الدالة أنسب للاستخدام في المستخلصات لأن لغتها الطبيعية أكثر إحكامًا من لغة العناوين .

ويجب أن نشير إلى أن هناك بعض التوجهات التي تشكك في جدوى استخدام كلمات التوقف ، أو ترى عدم قابليتها للتطبيق في اللغة العربية. وبرغم قلة التجارب ونقص الدراسات الموثقة حول هذه المسألة إلا أن مرد ذلك يعود إلى أسباب كثيرة منها الصعوبات الدلالية بالإضافة إلى عوامل أخرى مثل عدم وضوح الهدف وطبيعة المعلومات وكذلك الإمكانات البشرية وخصائص النظم المتاحة لاسترجاع المعلومات في المكتبات .

بيد أن فوائد استخدام كلمات التوقف علاوة على أهميتها في مجال الاسترجاع ، فهي تخفف العبء على النظم والاجهزة التي تتعامل مع كمية هائلة من المعلومات والنصوص . كما أن اللغة العربية تمتاز بالثراء المعجمي الذي يعني أن الكلمات غير الدالة متوفرة بكثرة في اللغة الطبيعية . ومن ناحية الدلالة والاسترجاع فإن كلمات التوقف قد استخدمت منذ أمد في كشافات نصوص القرآن الكريم مما يؤكد الأهمية لاستخدامها في نظم الاسترجاع بالمكتبات .

خصائص كلمات التوقف :

رغم صعوبة تحديد كلمات التوقف العربية بشكل دقيق وقاطع إلا أنه يمكن القول بأن كلمات التوقف مفردات غير دالة على مفهوم أو موضوع محدد وتمتاز بالخصائص التالية : -

١ - إنها مفردات لا تحمل مضامين أو معاني دلالية مستقلة إذا تم فصلها عن السياق .

٢ - إنها كثيرة الورود في العناوين والنصوص العربية .

٣ - إنها حروف أو أدوات لازمة لتركيب الكلام العربي والربط بين مفرداته وأجزائه .

٤ - إنها قد تكون صفات ونعوتا لمفاهيم أخرى .

٥ - إنها كلمات عامة وغير مخصصة بنوع من الموضوعات أو حقل من العلوم .

٦ - إنها مفردات لا يحتمل السؤال عنها وتذكرها عند البحث والاسترجاع.

٧ - ومن الناحية التركيبية لا يمكن تكوين جملة مفيدة ولها دلالة موضوعية عند استخدام كلمات التوقف لوحدها مثل :

نبذة عن أحسن الإرشادات لاستخدام أنواع النظريات والتأملات حول أوضاع المسائل والأمور المساعدة لأغراض تفاصيل الأدوات من زاوية الممارسة العامة في المجال المحتمل ودور الخفايا المشكلة من مضامين وزوايا ذات أبعاد وأشكال مقدرة بالنسبة لأسباب وأحوال الاتجاهات والمواقف والمؤشرات الواضحة حسب ماهية النشاطات المختارة وتطورات المفاهيم...

وعند تتبع العناوين العربية نجد أنها قلما تخلو من حروف أو مفردات حشو يستخدمها المؤلفون باللغة الطبيعية ، وهي لا تضيف معلومات دلالية مهمة في العناوين مثل العناوين التالية :

(في) النحو والصرف

(حول) الثقافة العربية

(دراسة حول) التنمية الإدارية

كما يستخدم المؤلفون الكثير من حروف العطف والجر التي ترد بحكم الضرورة اللغوية للتعبير عن الربط بين مفهومين متلازمين سواء كان الربط بين موضوع وآخر أو بين موضوع ومكان مثل (الإدارة والتنمية في الخليج العربى) بل إن العنوان الواحد قد يتضمن عشرات من الأدوات النحوية ما بين حروف جر وعطف وأسماء موصول أو ظروف أو أسماء خمسة . من هنا نجد أن كلمات التوقف العربية ، إما أنها حروف وأدوات نحوية وإما مفردات قد تدل على معانٍ معجمية أو مفاهيم عامة لكنها لا تحمل مضامين اصطلاحية في أي علم من العلوم . ومن ثم فإن التخلص من خلال تباديل الكشافات مهم من ناحية الاسترجاع وإلا فإن الكشاف سوف يكتظ بعدد لا حصر له من الكلمات الفارغة التي تعيق استرجاع المعلومات من عدة أوجه .

وتشمل كلمات التوقف العربية الأدوات النحوية من الحروف والروابط مثل (أل) التعريف ، وحروف الجر المنفصلة والمتصلة ، وحروف العطف والاستفهام والنفي والتعجب والنداء ، وكذلك الظروف الزمانية والمكانية وأفعل التفضيل كما تشمل جميع الضمائر وأسماء الإشارة والموصولات والأسماء الخمسة ، وبعض الأعداد والمضافات والأفعال ، وقد تكون كلمات التوقف منفصلة أو ملتصقة بأوائل الكلمات أو ملحقة في أواخرها .

كما تشمل كلمات التوقف بعض الكلمات ذات المفاهيم العامة والمزدوجة الدلالة ، وهي الكلمات التي تحمل مفهومين أحدهما اصطلاحي والآخر معجمي أو عام ، ويمكن اعتبارها من كلمات التوقف إذا كانت ترد في الغالب بمضامين عامة وغير دلالية . وهناك مفردات أخرى لا يمكن الجزم بتأكيد دلالاتها إلا

بالتعرف عليها ضمن سياقاتها . ولذا لا يمكن اعتبارها من الكلمات الدالة اعتماداً على طبيعة النص والتخصصات التي تهم المكتبة أو نظام المعلومات .

ومن حيث التقسيم الصرفي فإن كلمات التوقف ترد في صيغ أسماء ومشتقات ثم حروف ثم أفعال ، ولكن الأسماء والمصادر المشتقة لا تكون إلا كلمات دالة في الغالب ، وهي إن جاءت غير دالة فإما بسبب أنها عامة جدا ، أو مزدوجة المعنى ، أو بسبب أنها دالة على مفهوم شكلي لأسلوب معالجة المعلومات مثل :

(أدوات ، إيضاحات ، بحوث ، تقارير ، تجارب ، دراسة ، دليل ، رسالة ، قسم ، قضية ، كتاب ، مرجع ، أعمال) ونحو ذلك . والمفاهيم الشكلية لها معالجة خاصة باللغات الاصطناعية وفي بعض نظم المعلومات .

أما معايير تقييم الدلالة وأهمية الكلمات فهي ببساطة مسألة لا تتأثر بالذائقة الأدبية أو اللغوية ، وإنما تعتمد على المعرفة الموضوعية باهتمامات المستفيدين وبلغة المعلومات المستخدمة في نظام المعلومات . وهذا يعني معرفة المصطلحات المستخدمة في الإنتاج الفكري والحقول العلمية التي يخدمها نظام المعلومات . هذا إلى جانب الاستناد على المكانز وقوائم رؤوس الموضوعات وكذلك الموسوعات والمعاجم المتخصصة . كما أن تقدير أهمية دلالات الكلمات في الاسترجاع مسألة نسبية ، فالكلمات قد تتفاوت دلالاتها حسب اختلاف الحقول واهتمامات نظم المعلومات وأغراض المستفيدين . ففي حقل اللغة العربية تصبح الحروف والأدوات النحوية كلمات دالة وموضوعات مهمة في الاسترجاع مثل العناوين التالية :

لعل في القرآن الكريم

رسائل المباحث المتعلقة بـ (من) الشرطية

فعندما تصبح المفردات غير الدالة مثل (لعل ، من) محورا للبحث والدراسة، فمن الطبيعي أن يكون لها مداخل في نظام الاسترجاع ، وبالتالي تعد من الكلمات الدالة ، وهذا يعني أن نوعية كلمات التوقف تختلف حسب الموضوعات والمستفيدين وأغراضهم ونوعية المعلومات أو النصوص المختزنة في نظام الاسترجاع .

أنواع كلمات التوقف :

تمتاز اللغة العربية بالمرونة وكثرة التصاريف التي قد يصعب أو يستحيل إيجاد قائمة توقف ثابتة في كل الأحوال والسياقات باللغة الطبيعية، خصوصا إذا كان نظام الاسترجاع في مكتبة متخصصة في الحقول الاجتماعية المتداخلة . ولكن بسبب قلة التجارب العربية الموثقة فيما يتعلق بتصميم واستخدام الكشافات واسترجاع المعلومات باللغة العربية الطبيعية لا بد من حصر كلمات التوقف الممكنة بالاعتماد على دراسة إحصائية ولغوية لمعرفة أنواع ودلالات الكلمات في عناوين المطبوعات العربية التي تضمها فهارس المكتبة أو الكشافات المطلوب تحسيبها وذلك يستلزم تعرف طبيعة صياغات الكلمات وأنواعها واستخداماتها في العناوين ومعرفة سكوناتها من المفردات غير المهمة التي تعيق استرجاع المعلومات .

وفي سبيل إيجاد قائمة توقف الكلمات العربية ، ورغبة في فرز ودراسة الكلمات غير الدالة والشائعة في عناوين الكتب والمقالات ، فقد تم القيام بمسح

إحصائي تجريبي لما مجموعه (١٨٠٠) عنوان من الكتب ومقالات الدوريات في مختلف حقول المعرفة الرئيسية بما يشمل حقول الدين والفلسفة وعلم النفس والاجتماع والتربية والإدارة والقانون والاقتصاد والتاريخ والدراسات الأدبية ، وفي بعض الكتب العامة في العلوم البحثية والتطبيقية . ولم يدخل في عملية البحث الأعمال الإبداعية من شعر ورواية أو نثر فني ونحوها ، لأن مكونات عناوينها إما غامضة أو رمزية ، فهي في الغالب لا تحمل مضامين موضوعية يمكن الاستناد عليها عند استرجاع المعلومات . فدلالة الكلمات في العناوين الرمزية والمجازية لا تطابق المفاهيم أو المعلومات التي تعالجها هذه الوثائق أو الأعمال الإبداعية والنصوص الأدبية . وهذا ما يؤكد أن فعالية استخدام كشافات التباديل باللغة الطبيعية تتحقق في استرجاع المعلومات الحقائقية أو العلمية كما في نظم المكتبات المتخصصة ومراكز المعلومات .

ويهدف مسح العناوين بالدرجة الأولى إلى تكوين قائمة تجريبية أساسية للكلمات من الحروف والمفردات غير الدالة والتي يكثر استخدامها في العناوين واللغة الطبيعية . ونتيجة لمسح العناوين وتسهيلا لمعالجة كلمات التوقف فقد تم تقسيمها إلى نوعين : أحدهما الحروف والأدوات النحوية والربط بين المفاهيم، والنوع الآخر الكلمات غير الدالة الأخرى وهي المفردات غير الدالة على مفاهيم ثابتة في حقل من حقول المعرفة الموضوعية .

الحروف والأدوات :

تكتظ العناوين العربية بأشكال متعددة من الروابط والأدوات النحوية والحروف التي تشمل حروف الجر والعطف ، وكذلك الظروف الزمانية والمكانية

والضمائر ، وأسماء الإشارة وأدوات الاستفهام وأسماء الموصول والأسماء الخمسة سواء وردت هذه الأدوات منفصلة أو متصلة بغيرها من الأسماء والضمائر مثل (في ، فيه ، فيما ، عليه) أو جاءت في صيغ إعرابية معينة مثل (ذي) من الأسماء الخمسة في حالة الجر لأنها في الغالب تقع مضافة لما بعدها في العناوين . كما تأتي هذه الأدوات أحادية الحروف مثل واو العطف والياء والباء واللام والكاف من حروف الجر ، أو جاءت ثنائية مثل (من ، عن ، في) من حروف الجر وكذلك (لو) من حروف العطف و (هل ، كم ، ما) من أدوات الاستفهام ، أو كانت على ثلاثة أحرف مثل حرف الجر (إلى) والظرف (عند) أو من أدوات الاستفهام الثلاثية مثل (أين ، كيف) أو كانت رباعية مثل (ماذا ، هكذا) (النحاس) .

ولقد بلغ مجموع الأسماء والحروف والأدوات النحوية التي تم حصرها (٥١) أداة نحوية ، منها (١٣) حرفا وأداة تكررت أكثر من تسع مرات بلغ مجموع تكرارها (١١٠٤) مرة في الألف وثمانمئة عنوان ، سواء وردت في العنوان الواحد مرة واحدة أو أكثر من ذلك . أما الحروف والأدوات التي وردت أقل من عشر مرات فقد بلغ مجموعها (٣٨) أداة ظهرت في بعض العناوين الطويلة ، أو ذات الصياغات الأدبية والإرشادية ، وفي بعض عناوين كتب التاريخ أو المواد المترجمة .

ونورد فيما يلي جدولاً للحروف والأدوات وهي أكثر كلمات التوقف وروداً في العناوين مرتبة تنازلياً حسب نسب تكرارها .

أكثر الحروف ورودا في العناوين العربية	عدد المكرر	المفردات التي تكررت أقل من عشر مرات
في	٣٩٢	أم، أو، أي، أيها، أمام، التي،
(و) العطف	٣٦٨	بعض، لبعض، كل، غير، حول،
(ل)(لام الجر المتصلة بكلمة دالة)	١٠٧	نحو، ذات، ذى، لها، عليها،
من	٥٨	له، عليه، بعد، قبل، حتى،
(ب) (باء الجر المتصلة بكلمة دالة)	٣٢	كما، بما، فيما، لما، منه، فيه،
على	٢٦	هو، هي، لدى، هل، معك،
إلى	٢٣	ماذا، لماذا، هكذا، لمن، تحت،
بين	٢٠	منذ.
عند	١٨	
ما الاستفهامية	١٨	
عن	١٦	
مع	١٦	
كيف	١٠	
مجموع التكرار	١١٠٤	

وقد تم رصد ما ورد خلال الحصر والملاحظة للعناوين ، أما التصاريف والصيغ الإعرابية المحتملة لأي من المفردات المسرودة فلم يتم ذكرها هنا رغبة في عدم التدخل في معطيات المسح ، ثم إن الهدف هنا تكوين قائمة توقف أولية يمكن الاستفادة منها لحصر قائمة أشمل عند التفكير في إعداد قائمة توقف عامة وأساسية لتلائم مختلف الحقول في المكتبات المتخصصة .

وعند النظر في قائمة التوقف التجريبية نجدها لا تمثل إلا نسبة ضئيلة من الحروف والأدوات النحوية التي يمكن استخدامها في النصوص العربية ، والسبب في ذلك أنه لم يتم حصر إلا ما ورد خلال مسح العناوين . ونجد مثلا حرف الجر (في) قد يتصل بضمير المؤنث المفرد (فيها) وقد يرد في العناوين بهذه الصورة لكنه لم يرد خلال المسح التجريبي ولذا تم إغفاله في القائمة أعلاه.

وبما أن الكلمة في مفهوم الحاسوب هي الوحدة المؤلفة من حرف أو أكثر والمحاطة بفراغين ، فإن هذا المفهوم لا بد أن يطبق على أشكال الحروف والكلمات غير الدالة الأخرى على اختلاف تصاريفها واللواحق المتصلة بها كتابيًا . ومن هنا نجد أن الأدوات المسرودة تشمل المجردات من حروف وأدوات أخرى قد تلحقها . فإذا كانت اللواحق غير متصلة كتابيًا فإنها تجزأ عند حصر الكلمات ، لذا فقد اعتبرت (مالها ، ماله ، ما عليه ، ما عليها) كما يلي : (ما ، لها ، له ، عليه ، عليها) . وذلك لغرض معالجة هذه المفردات على أساس أنها وحدات مجزأة يحتمل ورودها بالشكل المنفصل في صياغات أخرى. أما إذا كانت ملتصقة أثناء الطباعة فإنها تعد كلمات مستقلة وكلا الاحتمالين وارد في اللغة الطبيعية ، على أن اعتبارها وحدات مدمجة في الطباعة قد يثير لبسًا حيث إن (ماله ، مالها) قد تعني كلمة (مال) المضافة إلى الضميرين .

إما إذا كانت الحروف متصلة بغيرها من ضمائر ونحوها فإنها تعد كلمة واحدة مستقلة ومتعددة حسب اختلاف الزوائد واللواحق المتصلة بها مثل (منها ، منه ، بما ، لما ، كما ، فيما ، عليه ، عليها ، لبعض ، لماذا) ونحو ذلك. كما

تم التفريق بين حروف الجر المتصلة مثل اللام والباء في حالة اتصالهما بكلمات دالة ، عن حالة اتصالهما بحروف أو أدوات نحوية غير دالة . فالحروف المتصلة بكلمات غير دالة تعد وحدة مستقلة ضمن كلمات التوقف مثل : (أيها ، منه ، فيه ، معك) ونحوها من الأمثلة السابقة . أما الحروف المتصلة بكلمات دالة مثل (بالأدب ، للأدب) ونحوها فسيتم معالجتها فيما بعد .

وبمراجعة جدول كلمات التوقف المحصورة نجد أن نمط الترتيب الاستخدامي للحروف التي وردت في المسح الاستطلاعي للعناوين يكاد يتطابق مع النتائج التي توصلت إليها دراسة (الخولي) وكذلك بحث (داود عطية عبده) المشار إليهما سابقاً على الرغم من اختلاف نوعية العينات . فعينة الدراسة هنا هي العناوين ، بينما انصبت دراسة (الخولي) وبحث (عبده) على النصوص . وخلاصة هذه الدراسات تدل على أن حرف الجر (في) وكذلك واو العطف هما أكثر الحروف العربية وقوعا في العناوين ، وفي النصوص المكتوبة الأخرى ، يليهما في الترتيب حروف أخرى منفصلة أو متصلة ذات تكرار مرتفع في اللغة العربية . وهذا يدل على أن نمط تركيب الجمل العربية موحد ما دامت النصوص حرة وطبيعية .

وأخيراً فلا تشمل كلمات التوقف ما ذكر هنا فحسب ، إذ أن مراجعة سريعة لبعض كتب اللغة والنحو تدل على أن هناك المئات من الحروف والأدوات والروابط النحوية الكثيرة التي يمكن مراجعتها عند البدء في تكوين مفردات توقف ، على أن يكون المعيار الأساس في الاختيار هو مدى تكرار وشيوع الكلمات في عناوين المطبوعات في المكتبة .

أما إذا كانت الأدوات من أسماء وحروف أو أفعال نادرة الوقوع في العناوين فليس هناك حاجة لحصرها . ونجد أن بعض الحروف والأدوات النحوية من

النواصب والجوازم وأسماء الأفعال والأفعال الجامدة وغيرها من حروف المعاني نادرا ما تقع في النصوص الحديثة فما بالك بعناوين الكتب والبحوث العلمية ، والأمثلة على ذلك كثيرة مثل (إذن ، لا النافية ، أن الناصبة ، ليت ، أضحى ، إلا ، قد ، كأن ، لعل ، كاد ، مهما ، لم ، لن ، هيهات) ونحو ذلك .

والخلاصة : أنه يمكن الجزم بثقة عالية بأن الحروف والروابط النحوية الشائعة قلما تأتي على أساس أن لها دلالات موضوعية ، ولذا يمكن إدراجها ضمن كلمات التوقف . أما مشكلاتها فهي تركيبية أكثر منها دلالية ، وهذا ما يؤثر على الاسترجاع أحياناً في استخدام اللغة الطبيعية حسب طبيعة النصوص أو المعلومات المستهدفة كما سنرى عند دراسة صياغات العناوين .

المفردات غير الدالة :

علاوة على الحروف والروابط النحوية هناك الكلمات ذات المفاهيم غير الموضوعية . والمفردات غير الدالة هي الكلمات العامة أو اللا واصفة والتى لا تدل على معنى اصطلاحي محددأو قائم بذاته ، أو قد تكون ذات مفهوم عام أو شكل أو كلمات وصفية مرتبطة بوجهات نظر ذاتية وغير موضوعية . وعند فحص العناوين العربية نجدها تتضمن الكثير من الكلمات الفارغة أو التي لا تحمل مفاهيم مستقلة بذاتها . ورغم إمكانية جمع قائمة طويلة من هذا النوع إلا أن هناك محاذير أغلبها دلالية أو تركيبية تتعلق بالسياق والمفاهيم الأخرى المحيطة بكلمات التوقف المنتخبة .

ويجب أن نشير إلى أن هناك صعوبات جمة تكتنف الخوض في هذا النوع من كلمات التوقف . فإذا كان من الممكن وبسهولة تمييز الحروف والأدوات النحوية فإنه قد يستحيل في ظل التجارب الأولية إيجاد قائمة توقف جامعة مانعة

لمختلف التغيرات الدلالية في عناوين الكتب والمقالات العربية . إلا أن ذلك لا يحول دون معالجة بعض الصعوبات الدلالية والتركيبية التي تحيط بعملية التمييز بين الكلمات الدالة والأخرى غير الدالة مع إبراز الفروق في سياقات أمثلة تطبيقية من واقع العناوين العربية أو اللغة الطبيعية .

كما أن تحديد الكلمة الدالة من غير الدالة يعتمد على طبيعة النصوص وعلى أغراض التكشيف . فإذا كان النص من (جملة أو عبارة أو آية أو حديث أو مثل أو حكمة أو بيت شعر) هو المستهدف من عملية الاسترجاع ، فإن التكشيف قد لا يعتمد على استخدام مصطلحات محددة وإنما قد تكون مداخل الكشاف كافة الألفاظ الموجودة في سياق النص . أما إذا كان المستهدف من الاسترجاع هي المعلومات فإن مداخل التكشيف لا بد أن تعتمد على مصطلحات وكلمات دالة على مفاهيم محددة من الكيانات أو العمليات أو المجردات ونحو ذلك من موضوعات قابلة للاسترجاع .

ولعل من أشد الصعوبات التي تواجه تجميع قائمة التوقف أن دلالات الكلمات متغيرة حسب السياقات المحيطة بها . وذلك بسبب أن الكلمة مزدوجة الدلالة . وبموجب ذلك قد تكون الكلمة مفتاحية ومهمة في الاسترجاع أحيانا ، إلا أنها ترد في سياق آخر بلا قيمة دلالية مهمة في الاسترجاع . ونورد فيما يلي بعضًا من هذه الكلمات في سياقين : أولهما يظهر بأن الكلمة دالة ، والآخر يبين أن الكلمة غير دالة أو مهمة في استرجاع المعلومات :

إعداد البحوث العلمية

بحوث مؤتمر الحاسب الآلي

الأحوال الشخصية

الأحوال العامة في حكم المماليك

إدارة الأعمال

أعمال نجيب محفوظ

نظرية التطور البشري

تطور الاقتصاد السعودي

القطاع العام والخاص

قطاع الزراعة

النحو والصرف

نحو إدارة علمية

صناعة الكتاب العربي

كتاب التفسير

ولذا سيكون هناك صعوبة في حصر الكلمات العامة غير الدالة ، وبالذات إذا كانت ترد مزدوجة الدلالة بصورة متساوية في عناوين مكتبة معينة ، لذا يستحسن حسابها على أنها كلمات دالة حتى ولو أربكت الاسترجاع ، ولا شك أن هناك الكثير من الكلمات التي ترد في الغالب بلا دلالة وترد نادرا بمضامين مهمة مثل كلمة (مقدمة) في «مقدمة ابن خلدون» . فهذه الكلمة وما يشابهها لا شك أنها لا واصفة أو مهمة في الاسترجاع الموضوعي ، ولذا توضع ضمن قائمة التوقف مع استثنائها من مقدمة ابن خلدون .

كما يستخدم المؤلفون في عناوين كتبهم أو مقالاتهم كلمات إنشائية وحشوا على الرغم من أن الموضوعات التي يعالجونها قد تكون علمية . ومن هذه الكلمات (أهمية ، ضرورة ، حتمية ، الأسلوب السليم ، الأسلوب الأمثل ،

أفضل الأساليب ، المنهاج الواضح) ونحو ذلك من عبارات كما في العناوين التالية :

* الأسلوب الصحيح للطباعة على الآلة الكاتبة العربية

* إدارة التنمية محور أساس من محاور التنمية الشاملة

* العينات كأحد الفروع المهمة في علم الإحصاء

فصياغات هذه العناوين تتضمن مفردات لا واصفة وغير مهمة في الاسترجاع ، ولذا يمكن وضعها ضمن قائمة التوقف لتشمل الكلمات التالية (الأسلوب ، الصحيح ، محور ، أساس ، محاور ، شاملة ، كأحد ، النوع ، المهمة) إلى جانب حروف الجر ، على أن ذلك قد يهون أمام استخدام بعض المؤلفين لكلمات مجازية جوفاء في عناوين الموضوعات المتخصصة مثل :

* ينابيع التطور في إدارة موارد النمو

* الاحتمالات مبشرة : دورة الإدارة في مرحلة الانفتاح

فالمفردات غير الدالة في المقالين السابقين هي (ينابيع ، احتمالات ، مبشرة، دورة ، مرحلة ، في) مع أن بعض هذه المفردات قد يأتي بدلالات موضوعية في سياقات وحقول أخرى .

ومن الأساليب الترويجية التي تستخدم في صياغة العناوين تلك التي تسعى إلى جذب القارىء وتلمس احتياجاته المعلوماتية والنفسية مثل :

* القواعد الذهبية لاتقان العربية

* النور الساطع في قواعد اللغة العربية

* تعليم اللغة الانجليزية في خمسة أيام

* أسرار وعجائب الشفاء بالحبة السوداء

* سري وعاجل للنساء فقط : أسرار المرأة وحياتها

* عزيزي مريض الروماتيزم هذا هو الطريق إلى الشفاء

* تحفة المستفيد بتاريخ الأحساء

بينما نجد أن الكلمات الدالة في هذه العناوين هي (اللغة العربية ، الانجليزية ، الحبة السوداء ، المرأة ، النساء ، الروماتيزم ، الأحساء) وما عداها فهو حشو .

والواقع أن بعض المؤلفين لا يضعون في حسبانهم تقنيات الفهرسة والاسترجاع عند صياغة عناوين كتبهم أو مقالاتهم ، سواء من حيث التركيب اللغوي ، أو دلالة الألفاظ المعبرة عن المضمون الحقيقي للمادة المنشورة. فأحيانا يبدأ العنوان الرئيس بكلمات غير دالة ، وتكون الكلمات المفتاحية في آخر العنوان ، أو في العنوان الفرعي ، بينما يفترض العكس لان الجزء الثاني عنوان شارح لما قبله مثل :

زمن لكل الأزمنة : نظرات وآراء في الفن

بين القديم والجديد : دراسات في الأدب والنقد

وكثيرا ما تأتي عناوين بعض الكتب والبحوث على هيئة مستخلص دال على محتوى المادة مما يؤدي إلى طول العنوان ، وضرورة استخدام الكثير من الروابط والمفردات غير الدالة إلى جانب تعدد الصيغ الصرفية ، ورغم أن الطول المعقول للعنوان قد يكون ميزة محمودة ، إلا أن ذلك يعتمد على مهارة المؤلف في الصياغة وإظهار الكلمات الدالة دون أن تطغى الكلمات غير الدالة على

سياق العنوان . كما أن الطول المفرط للعنوان قد يثير مشكلات عند التباديل والاسترجاع في بعض النظم وذلك يعتمد على طول حقل العنوان وعدد الكلمات أو الحروف التى يمكن تباديلها واسترجاعها في السياق ، وهل ذلك يشمل العناوين الفرعية أم لا ؟ ومن العناوين الطويلة ما يلي :

* إجراءات المراجعة للشركات التي تساهم الدولة في رأس مالها أو تضمن لها حدا أدني من الأرباح .

* الجديد في قوانين التموين والتسعير الجبري : أحدث التشريعات والقرارات التموينية والقيود والأوصاف معلقا عليها بأحكام القضاء وآراء الفقهاء.

كما نجد عناوين الكتب تبين بعض توجهات التأليف التي تنحو إلى العموميات وتأليف الكتب الدراسية ، ويتضح ذلك في آلاف العناوين التي تبدأ بكلمات غير دالة مثل (مبادىء ، أصول ، مقدمة ، مدخل ، أسس ، أساسيات ، محاضرات ، قراءات) ونحو ذلك من المفردات غير الدالة المتكررة بصورة عالية في العناوين .

أما عناوين كتب التراث القديمة فيغلب عليها السجع الذي استدعته الحاجة لحفظ وتذكر العناوين ، مما استلزم استخدام الكثير من الكلمات الإيقاعية الفارغة من الدلالة . مثل كتاب ابن القيم الجوزي (الجواب الكافي لمن سأل عن الدواء الشافي) فالكلمة المفتاحية في العنوان هي (الدواء) . كما قد تصاغ العناوين بمنهج إيحائي غامض ؛ بل لا يمكن تفسير دلالة نص العنوان إلا بالاعتماد على السياق الخارجي والمعرفة السابقة الخارجة عن النص كما يقول اللغويون . ومثال ذلك (الدر الثمين والمورد المعين) للمالكي وعنوان (الدر المنظوم لذوي العقول والفهوم) لعبدالله الحداد .

كما يكثر في عناوين كتب التراث العربية استخدام كلمات حشو شائعة مثل (كتاب ، رسالة ، حاشية ، أخبار ، متن ، المنتخب ، المنتقى ، المغني) ونحو ذلك من الكلمات التي لا تنفع في استرجاع المعلومات باللغة الطبيعية بالاعتماد على دلالات الكلمات وحدها ، وبالذات إذا كان العنوان الرئيس غامضًا ولم يكن مشروحا بعنوان فرعي قابل للاسترجاع مثل بعض القواميس العربية القديمة (المرصع ، الكافي ، الصحاح ، العباب الزاخر واللباب الفاخر). على أن قواعد الفهرسة الوصفية تدعو إلى التخلص من بعض الكلمات مثل (كتاب) إذا وردت في أول العنوان ولم يؤثر حذفها على الصياغة مثل (كتاب آراء أهل المدينة الفاضلة) للفارابي ونحوه .

ومع ذلك فإن بعض الكلمات التي تبدو غير دالة على محتوى الكتاب إلا أن لها أهمية خاصة في استرجاع الكتب المرجعية المشهورة مثل عناوين القواميس المشهورة أو كتب التراث العربي خصوصا في المكتبات التي تعود الباحثون فيها على حفظ العناوين المرجعية المشهورة مثل (الأغاني ، تاج العروس ، يتيمة الدهر ، مغني اللبيب ، رياض الصالحين) وغيرها من الكتب والمراجع التي تستلزم استثناء بعض الكلمات غير الدالة لأنها مهمة في استرجاع عناوين معينة ، غير أن كشافات التباديل بالاعتماد على الكلمات الدالة تستجيب أكثر للاحتياجات البحثية التي تتجاوز مجرد استرجاع عناوين محددة ومعروفة ، أو مقصودة بذاتها من قبل القراء .

ومن ناحية التركيب فإن كلمات التوقف تأتي مفردة أو متصلة بضمائر ولذا توضع ضمن قائمة التوقف حسب التصاريف المحتملة للضمائر وبالأخص للمفرد المذكر والمؤنث ، وذلك على أساس أن هذه المفردات مرتبطة لغويًا

وسياقاً بكلمات أخرى ظاهرة وأقوى دلالة في تفسير نص العنوان والمحتوى الموضوعي للوثيقة مثل :

اتجاهاته	حاضرها	استخداماتها
أحواله	مضامينها	أعماله
الآخذة	المتعلقة	ألقاها
أهميتها	مفهومه	أوصافه
تطوراته	آراؤه	ماضيها
معاييره	مواقفه	نشاطاته
مستقبله	ماضيها	وضعها

فاتصال الكلمات غير الدالة بضمائر أكثر من اتصال الضمائر بالكلمات الدالة في العناوين التالية :

الفعالية الإدارية : أهميتها ، مداخلها ، مقاييسها ومشكلاتها

البحث في الإدارة العامة : نشأته ، مشاكله ، معوقاته

الأسواق المالية : أهدافها ، أنواعها وفاعليتها

أما اتصال الضمائر بالكلمات الدالة فقليل مثل الكلمات التالية (استثماراته ، اقتصادياته ، ثقافته ، متاحفها ، مدارسها) مثل : (تطوير دور الغرف التجارية في تشجيع القطاع الخاص على توجيه استثماراته في النشاط الصناعي) .

كما تأتي العناوين على شكل أسئلة تبدأ بأدوات الاستفهام وأفعال مختلفة مثل :

كيف انتشر الإسلام ؟	مؤيد الكيلاني
كيف تعد بحثا أو رسالة ؟	أحمد شلبي
لماذا تأخر المسلمون ؟	شكيب أرسلان
ما الدبلوماسية ؟	عز الدين فودة
ما هو علم المنطق ؟	يحيى هويدي

كما قد تكون أدوات الاستفهام في موضع آخر غير بداية العنوان ، مثل المقال التالي (التخطيط للقوى العاملة . كيف ولماذا؟) . ولا شك أن حصر أدوات الاستفهام الشائعة ضمن قائمة التوقف يدعو إلى ضرورة التفكير في استبعاد الأفعال والضمائر التي تليها مع مراعاة صياغة العنوان ومكوناته من الكلمات الأخرى ، سواء كانت دالة أم غير دالة .

ومن الأفعال غير الدالة التي يكثر وقوعها في العناوين لا سيما في صيغ الاستفهام ، سواء جاءت في صيغ المضارع للمذكر والمؤنث وهي الأكثر ، أو جاءت في صيغ الماضي مثل :

يتعلق	يعمل	يتم
تتعلق	تعمل	يمكن
تعنى	يتكامل	أمكن
يعنى	تتكامل	تعرف
كان	يطبق	تعلم
يكون	تطبق	يجري
تكون	تتم	يريد
نستخدم	نستعمل	

ومثل ذلك يقال بالنسبة لأفعل التفضيل الذي يأتي بلا دلالة في الغالب كما في العناوين التالية :

أحلى عشرين قصيدة

أحسن القصص

أوضح الأساليب في الترجمة والتعريب

أفضل سنوات المرأة

كما يرد في العناوين الكثير من الأعداد الرقمية والحرفية ومميزاتها ، فيمكن جعل الأعداد من كلمات التوقف والإبقاء على تمييز العدد إذا كان مهمًا في الاسترجاع مثل (٢٠ قصة قصيرة) أو الإبقاء على العدد والتمييز كما في بعض كتب الأدب الإبداعي مثل (ألف ليلة وليلة) . ولكن التمييز قد لا يكون ذا دلالة مهمة في كثير من المجالات الموضوعية مثل :

أربعون درسًا وتمرينًا في قواعد اللغة العربية

مائتان وخمسون سؤالا وجوابا في تربية الحيوان والأسماك

مائة وخمسون ألف سؤال عن أمراض الأطفال

مائة سؤال عن الإعلام

فالكلمات الدالة في هذه العناوين هي (اللغة العربية ، الحيوان ، الأسماك، أمراض ، أطفال ، إعلام) كما يكثر استخدام الأعداد في عناوين الكتب التاريخية وفي وقائع المؤتمرات والاجتماعات مثل :

تاريخ أوربا في القرن الثامن عشر

العالم سنة ٢٠٠٠

المؤتمر السابع للحاسب الآلي

والواقع أن أرقام التواريخ وأسماء الأشهر وأيام الأسبوع في العناوين قد لا تكون مهمة في الاسترجاع باستخدام اللغة الطبيعية ، لأن الأرقام قد تحمل معاني ودلالات متفاوتة كما أنه لا يحتمل البحث والاسترجاع بها في أغلب المكتبات إلا في بعض الوقائع والأحداث والأيام المشهورة على نطاق محلي أو عام مثل :

حرب ستة أكتوبر

الحرب العالمية الأولى

الحرب العالمية الثانية

والأرقام بحد ذاتها قد تكون مهمة في بعض الحقول العلمية إذا كانت رموزاً دالة على مفاهيم محددة أو أسماء رمزية كما في بعض المعادلات الرياضية والكيميائية وبرامج الحاسوب ورموز المشاريع والبرامج المشهورة التي قد تغطيها قواعد المعلومات المتخصصة .

ومن المشكلات التي تواجه استخدام كلمات التوقف وتحديد الكلمات غير الدالة ، أن بعض العناوين في الآداب والإنسانيات تتألف من مفردات غير دالة. ونجد أن الحروف والأدوات النحوية المؤكد عدم دلالاتها تأتي في حالات نادرة مهمة ليس من حيث تضمنها مفاهيم دلالية مستقلة ، ولكن من حيث قيمتها الاسترجاعية خصوصا في المكتبات التي تضم موضوعات متعددة ومتداخلة . فبعض المؤلفين والأدباء يصوغون عناوين كتبهم من مفردات غير دالة بل غامضة إلا لمن عرف محتويات الكتب مسبقا مثل :

بين بين	طه حسين
من قبل ومن بعد	طه حسين
عن هذا وذاك	غازي القصيبي

فنصوص العناوين والسياقات لا تدل على مفاهيم واضحة أو حتى عامة من حيث أن كل المفردات السابقة يفترض فيها أن تكون كلمات توقف لا قيمة دلالية أو موضوعية لها ، إلا إذا كانت المفردات ذاتها موضوعات كما في علوم النحو والصرف وفي المقالات والبحوث حول بعض الحروف والأدوات النحوية ومع ذلك فإن (بين ، من ، قبل ، بعد ، عن ، هذا ، و ، ذاك) تدرج ضمن قائمة التوقف ، مع التنبيه إلى احتمال ورودها في عناوين قليلة لا يمكن استرجاعها إلا بالاستثناء الفردي والمعالجة الخاصة أثناء الفهرسة كما في العناوين الثلاثة السابقة .

ولو نظرنا إلى المفاهيم المركبة من الناحية الدلالية لوجدنا أن بعضها يأتي من كلمتين مترابطتين ، إما على هيئة مضاف ومضاف إليه أو على هيئة صفة وموصوف مثل :

معالجة البيانات

التحليل الإحصائي

وعلى الرغم من أن كلمة (معالجة) غير دالة من حيث أنها عامة ، ولا تقف بدون تركيب مع مفهوم دلالي أخص ، إلا أنها جزء لا يتجزأ من المفهوم المركب والمخصص (معالجة البيانات) أو معالجة المرضى ومعالجة النفايات ، ومع ذلك يمكن القول بأن المضاف إما أن يكون كلمة دالة ، وإما أن يكون كلمة غير دالة ففي الحالة الأولى تعامل على أنها كلمة مفيدة في التحكم بالاسترجاع حتى ولو تكررت في سياقات مختلفة لمفاهيم متنوعة مثل (معالجة) وكذلك :

سباق الخيل

سباق التسلح

سباق السيارات

سباق الهجن

أما إذا كان المضاف كلمة غير دالة ولا تستخدم في تركيب المفاهيم فيمكن ضمها ضمن قائمة التوقف غير الدالة مثل :

أنواع الضرائب

استخدام الحاسب الآلي

أوضاع العمالة الأجنبية

واقع المدارس الحكومية

أسباب التخلف الاقتصادي

أما كلمة (تحليل) التي وردت في المثال الثاني فهي كلمة غير دالة أيضا لأنها عملية لمفهوم آخر متعدد ومتغير ، فهي غير مستقلة بذاتها في الغالب ، إلا أن حذفها سيترك الصفة بدون موصوف . وبذا يختل تركيب المفهوم ، فكلمة (الإحصائي) وإن كانت مفهوما دلاليًا مستقلا ، إلا أنها في هذا المثال تابعة لغويا لمفهوم مرتبط بها ، وعلى الرغم من أن الموصوف أهم من الصفة لغويا ودلاليا كما في (الإحصاء التحليلي ، المحاسبة التحليلية) إلا أنه ليس كذلك عندما تكون الصفة نسبة مشتقة من موضوع أو مفهوم دال ومخصصة لمفهوم عام .

ولذا يتم معالجة الصفة والموصوف على أنهما مفهوم مركب سواء من حيث دلالة المفردات أو من حيث التباديل والاسترجاع . وذلك على افتراض أن الباحث خلال الاسترجاع لن يستخدم كلمة (تحليل) لوحدها على أساس أنها كلمة بحثية وإنما سيتبعها بالصفة الدلالية الأقوى ، أو يربطها بمفهوم آخر

أخص خلال البحث وفق أسلوب الربط اللاحق باستخدام روابط الصلة المنطقية ويمكن أن نقيس على كلمة (تحليل) كلمات أخرى عامة إذا وردت موصوفة بمفاهيم دلالية . ومع ذلك نجد أن كلمتي (معالجة ، تحليل) وتصاريفهما تردان أحيانا دون تركيب وبلا دلالة مهمة في مواضع أخرى مثل :

معالجة التضخم المالي

التأخر الدراسي ومعالجة أسبابه

الأدب الحديث : تحليل ونقد

ويرى (فوسكت) بالنسبة للغة الإنجليزية أن الخواص النعتية للمفاهيم لا تقف وحدها ، وإنما يدل عليها بالأسماء . ويمكن أن نستخلص بأن الصفات والنعوت باللغة العربية تأتي من ناحية الدلالة على أربع حالات يمكن الاجتهاد في تمييزها على النحو التالي :

١ - أن تكون الصفات مشتقة من مفاهيم دلالية ، ولذا فهي مهمة في الاسترجاع مثل :

الإدارة المدرسية

الإجازات المرضية

الصناعة الفندقية

الصحة النفسية

التحليل الاقتصادي

التحليل الإداري

التحليل المالي

٢ – أن تـكـون صفـات عـامـة مخصصة للمفهـوم الموصوف ، وذلـك بـسبب أن المصطلح الدال مركب من صفة وموصوف ، ومن هنا فإن الصفة مخصصة للمفهوم ومهمة في التحكم بالاسترجاع مثل :

القانون الدولي الخاص

المكتبة العامة

القطاع الخاص

٣ – أن تكون الصفات لا تحمل معاني اصطلاحية مخصصة لمفهوم ، وإنما هي نعوت اجتهادية أو تفضيلية متغيرة ، ولذا فهي غير مهمة في الاسترجاع، ويمكن اعتبارها من كلمات التوقف غير الدالة مثل :

الكتاب المدرسي (الجيد)

طرق التدريس (القابلة للتطبيق)

الكتابة (الصحيحة)

المشاريع الحكومية (العملاقة)

ومع ذلك فإن هذا النوع من الصفات يرد في حالات قليلة ذات دلالات مخصصة لمفهوم اصطلاحي عام مثل :

الفنون الجميلة

الكتب النادرة

٤ – أن تكون الصفات غير مشتقة من مفاهيم دلالية وتابعة لموصوف غير دال ففي هذه الحالة نعتبر الصفة والموصوف من الكلمات غير الدالة مثل :

دراسة أولية

تقرير مبدئي

الأوضاع السائدة

مقدمة موجزة

الأحوال العامة

إطار نظري

فهذه الحالات وما شابهها تستوجب التفكير في كيفية معالجة كلمات التوقف إذا كانت مركبة مع مفاهيم أخرى ، أو كانت توابع غير دالة لمتبوعات غير دالة ، وهي أو جزء منها ضمن قائمة التوقف ، فالصفات غير الدالة واضحة في معظم الأحيان إلا أن كثرتها واختلاف تصاريفها للمفرد المذكر والمؤنث والجمع تجعل عملية الحصر صعبة . ولذا فهناك الكثير من المفردات غير الدلالية التي تحتاج معالجة خاصه أثناء إدخال البيانات والفهرسة .

وكما لاحظنا في استعراض بعض العناوين فإنها من وجهة النظر المعلوماتية تمثل بعضًا من المشكلات المتأصلة في نصوص وتركيب اللغة الطبيعية ، ويمكن تصنيف هذه المشكلات إلى ثلاث فئات :

١ - مشكلات في الدلالة :

ويتعلق ذلك بطبيعة الكلمات من حيث المضامين والمعاني الثابتة أو المتغيرة وصعوبة تحديد المفردات الدالة من غير الدالة أحيانًا .

٢ - مشكلات في التركيب :

وذلك بسبب تغير دلالات المفاهيم المركبة داخل السياقات المختلفة ، أو بسبب الزيادات واختلاف الإعراب الذي يغير دلالات أو شكل الكلمات .

٣ - مشكلات الاسترجاع :

ويعني ذلك الصعوبات التي تواجه الاسترجاع الموضوعي الفعال بسبب تغير الدلالات والتركيب الذي ينعكس على الترتيب والتباديل والوصول إلى المعلومات ، ومايستلزم ذلك من حلول لتطويع اللغة الطبيعية .

هذه الأمور تمثل تحدياً قويًا لخبراء المعلومات الذين يحاولون السيطرة على لغة المعلومات والتحكم بالاسترجاع عند تصميم نظم المعلومات التي تعتمد على اللغات الطبيعية كالعربية . فمع التراكم المعرفي وتشعب العلوم لم تعد عملية الاتصال بالمعلومات واسترجاعها تعتمد على المهارات الشخصية والذاكرة ، وإنما أصبحت تعتمد على مداخل نمطية ومصطلحات مقننة . وهذا مما يساعد على فاعلية استخدام كشافات الكلمات الدالة في مختلف النصوص، وبالذات في الحقول العلمية والمجالات التي تتصف بلغات محكمة ومصطلحات معيارية مستخدمة في مراحل إنتاج المعلومات ثم تنظيمها واسترجاعها .

تكوين قائمة التوقف :

عند البدء بتكوين قائمة توقف، لابد أن يكون الكشاف أو فهرس العنوان في المكتبة هو المنطلق الأساس لحصر وتجميع الكلمات غير الدالة، على أن يتم ذلك بطريقة منهجية. فيمكن عن طريق التتبع الألفبائي فحص أوائل العناوين وسياقاتها، ثم رصد المفردات غير الدالة باستخدام بطاقات مرتبة ألفبائيًا مما يسهل عملية المراجعة والإضافات والتعديلات ومعرفة نسبة تكرار الكلمات في عينة الدراسة.

ويمكن لأي نظام معلومات أو مكتبة أن تعد قائمة توقف مناسبة لتخصصاتها وطبيعة النصوص أو العناوين التي يتم تكشيفها واسترجاعها على أن يؤخذ في الحسبان أثناء حصر الكلمات جميع الصيغ الصرفية والإعرابية المستخدمة فعليًا في النصوص أو العناوين المختزنة . فالمفردات العربية تأتي على هيئة مفرد وجمع ومثنى ، والأفعال تأتي بالماضي والمضارع ، والصفات تكون للمذكر والمؤنث وللجمع سواء كان ذلك مرفوعاً أو منصوباً ، جمع مذكر سالمًا أو جمع مؤنث سالماً أو جمع تكسير . كما تأتي الكلمات مضافة إلى ضمائر متصلة كما تتصل الكلمات بحروف جر ، وغير ذلك من التصاريف المحتملة في لغة المعلومات مثل :

أثر ، أثره ، أثرها ، آثاره ، تؤثر ، تأثيره ، لتأثيره ، لأثره ، لأثرها ، مؤثراته ، تأثيرها ، لتأثيرها ، لمؤثراته ، تأثيرات ، لتأثيرات ، كتأثيرات ، مؤثرات ، المؤثرة .

أساس ، أسس ، أسسه ، أساسياته ، لأساس ، لأسس ، أساسي ، أساسيه، كأساس تطبق ، يطبق ، تطبيق ، بتطبيق ، لتطبيق ، تطبيقه ، بتطبيقه ، لتطبيقه ، تطبيقات ، تطبيقاته ، تطبيقها ، تطبيقاتها ، تطبيقاتهما ، تطبيقيا، بالتطبيق ، للتطبيق .

فهذه أمثلة لثلاث كلمات غير دالة ما عدا احتمال أن ترد كلمة (آثار) بمعنى العاديات القديمة ، إلا أن كثرة تصاريف هذه الكلمات تبين صعوبة السيطرة على تصاريف الكلمات العربية ، إلا بفحص مكونات العناوين أو

غيرها إذا كان مجال التكشيف في نصوص أخرى ، ثم رصد المفردات غير الدالة وتصاريفها المتكررة بكثرة ، آخذين في الحسبان إنه لا يمكن حصر جميع كلمات التوقف دفعة واحدة إلا في حالة تكشيف النصوص المحددة والثابتة مثل القرآن الكريم أو غيره من نصوص .

ويمكن إغفال بعض الكلمات وعدم إدخالها ضمن قائمة التوقف عند الجزم بندرة أو استحالة وقوعها في النصوص أو عناوين الكتب والمقالات التي تضمها قاعدة البيانات . وذلك يرجع إلى المجالات الموضوعية للوثائق والنصوص ، وإلى دلالات الكلمات المستثناة من قائمة التوقف مما يعتمد على نتائج الدراسة التي تعدها المكتبة . ولذا قد تظهر فروق بارزة بين مكونات قائمة توقف معدة لمكتبة طبية ، وبين قائمة أخرى معدة للاستخدام في مكتبة متخصصة في العلوم الشرعية أو القانونية أو الاقتصادية .

كما يجب الحذر في تحديد كلمات التوقف في الحقول الأدبية والإنسانية كما في المكتبات العامة والجامعية التي تتنوع فيها المعارف . وعند الشك في دلالات الكلمة أو تأرجحها بين الدلالة وفقد المضمون يستحسن ترك الكلمة على أنها دالة . ذلك أن ظهور بعض المفردات غير الدالة في التباديل والاسترجاع أهون من فقدان العناوين المطلوبة بالكلمات التي استخدمها الباحث مهما كانت مستويات الدلالة .

على أن هناك الكثير من كلمات التوقف التي يمكن الجزم بثقة عالية أنها لا يمكن أن تأتي دالة في أي سياق ، وهي الأدوات النحوية والحروف المنفصلة

والمتصلة مثل حروف العطف والجر وأل التعريف . لذا يمكن اعتبارها النواة الأولى لأي قائمة توقف في معظم نظم المعلومات الخاصة بالمكتبات .

وتتنوع قوائم التوقف حسب مشكلات الاسترجاع في لغة المعلومات المطلوب تطويعها لأغراض التباديل والاسترجاع ، ولذا يجب أن تحدد أغراض استخدام قوائم التوقف ونوعية المدخل ونقاط الوصول التي سيتم استخدام كلمات التوقف من أجلها مثل :

* كشافات العناوين

* المستخلصات

* كشافات الأسماء

* رؤوس الموضوعات

* أو نصوص أخرى محددة

ورغم تشابه تلك المداخل في نوعية مفردات التوقف على أساس أن نظام المعلومات واحد ، إلا أنه قد يوجد تفاوت طفيف في محتويات قوائم التوقف لكل مدخل حسب احتياجات المكتبة وطبيعة البحث والاسترجاع . فمثلا قد تكون الكلمات مثل (ابن ، بن ، آل ، أبو ، بنت ، با ، أم) ونحوها من كلمات التوقف المستخدمة في كشاف المؤلفين ، إلا أنها قد تكون كلها أو بعضها كلمات دالة في المستخلصات أو العناوين والنصوص الأخرى . أما فهارس الموضوعات التقليدية فقد لا تحتاج إلى قائمة توقف لأنها تعتمد على لغات اصطناعية محكمة وجميع المفاهيم مسبقة الربط ما عدا (أل) التعريف وبعض

الحروف أو الروابط القليلة سواء كانت متصلة أو منفصلة حيث يتم استبعادها لأغراض الترتيب والتباديل والاسترجاع أحيانا . أما إذا كان مجال استخدام كلمات التوقف لاسترجاع النصوص ، فإن تحديد كلمات التوقف يعتمد على عوامل مختلفة مثل أهداف نظام المعلومات وأغراض الاستخدام . فإذا كان النظام عامًا وغير محصور بفئة من الباحثين فإن كلمات التوقف قد تكون مشابهة لنظام الاسترجاع في المكتبات وبالذات فيما يتعلق بالحروف وأل التعريف .

ويمكن دائما التفكير في الاحتمالات والحلول المناسبة لجميع المتغيرات الدلالية والتركيبية التي قد تعوق عملية البحث والاسترجاع ، مع استغلال ميزان نظام المعلومات المستخدم من أجل تطويع لغة المعلومات . من هنا يجب أن نشير إلى أنه يمكن دائما إلغاء فاعلية بعض كلمات التوقف في عناوين محددة وقت الإدخال ، وبالذات العناوين التي لا يمكن استرجاعها عند فقد كلماتها غير الدالة . أما مشكلة الكلمات المتجانسة فيمكن للمفهرس معالجتها على حدة في كل عنوان ترد فيه أثناء الفهرسة خصوصا في النظم التي لا تفرق في الترتيب بين الكلمات المشكولة .

ويجب أن نشير إلى أن قائمة التوقف ما هي إلا أداة يسترشد بها المفهرسون أثناء الفهرسة وإعداد البيانات الببليوجرافية ، ولذا يستندون على القائمة الجاهزة في تحديد المفردات غير الدالة . أما إذا وردت كلمات غير دالة، وهي غير موجودة أصلا ضمن قائمة التوقف فإنه يرجع في اعتمادها على خبير الفهرسة الذي يكون مسئولاً عن الإضافات والتعديلات . وعند تحديث

القائمة بالحذف والإضافة على القائمة المختزنة في النظام فإن ذلك لن يؤثر على تباديل واسترجاع العناوين القديمة التي سبق إدخالها قبل التعديل . كما أن الكلمات المرغوب في إلغائها من قائمة التوقف ستظل تظهر في تباديل العناوين القديمة . كما يجب التنبه إلى الاتساق في استخدام كلمات التوقف بحيث لا يتفاوت استخدام الكلمات المتحدة دلالاتها مرة على أنها دالة ومرة على أنها كلمات توقف مهملة إلا في الحالات الاستثنائية الموجبة لذلك كما سبق إيضاحه . وهذا ما يدعو إلى ضرورة عمل التعديلات المناسبة في قاعدة بيانات المكتبة حتى تتوافق كشافات التباديل مع قائمة التوقف الأحدث والمستقرة .

وهكذا نرى أن تكوين ومعالجة قائمة التوقف عملية تتداخل فيها خبرات لغوية ومعلوماتية وحواسبية ، يشترك فيها المكتبي مع المبرمجين ؛ ولكن الحكم النهائي في اختيار كلمات التوقف ومعالجتها يرجع دائما إلى متخصصي المعلومات في المكتبة . كما أن التطبيق العملي هو المحك المناسب لمعالجة مشكلات الاسترجاع سواء كانت لغوية أم معلوماتية . وهناك بعض الحلول المنطقية للتخلص من بعض كلمات التوقف آلياً خلال التباديل والاسترجاع دون حاجة إلى حصر ودراسة المفردات ، وذلك بالنظر إلى السمات الغالبة والمشتركة للحروف والأدوات غير الدالة . فمثلا نجدها تأتي على حرف واحد منفصل مثل واو العطف المفصولة بفراغ عما بعدها ، كما تأتي ثنائية . فلو تم تصميم برنامج يستبعد من التباديل كل ثنائي تلقائيًا على افتراض أن معظم المفردات والحروف من هذا الصنف لوجدنا في اللغة العربية الكثير من الكلمات الثنائية

الدالة التي قد ترد بكثرة في العناوين مثل (أب ، أم ، فن ، حج ، جن ، حبّ ، رب ، طب ، حس ، شر ، دم ، نص) ونحو ذلك .

ونسرد فيما يلي بعض الكلمات ذات المفاهيم غير الدالة التي يكثر ورودها في العناوين العربية مع ملاحظة أنها ذكرت هنا بصيغة واحدة مع احتمال أن تكون دالة في حالات نادرة . وهذه القائمة لا تنطبق على كل احتياجات المكتبات المتخصصة ، كما أنها لا تشمل كلمات غير دلالية كثيرة ، إلا أنها مع الحروف قد تصلح نواة لقائمة توقف يمكن تحديثها بإضافة مفردات وتصاريف أخرى.

أبعاد	اتجاهات	أحسن	أحوال	أدواته	آراء
إرشادات	أسباب	استخدام	استعمال	أشكال	أغراض
آفاق	أفضل	أنماط	أنواع	أوضاع	أوجه
بواسطة	تأملات	تجارب	تصورات	تطورات	تعريف
تفاصيل	توجهات	جوانب	بالنسبة	جديد	جميع
حلول	خفايا	خلال	خواطر	دعائم	رؤية
زاوية	ضوء	طريقة	عام	عامة	عبر
عناصر	فروع	قضايا	كبير	كثير	كل
لمحات	مبسط	مجال	محتمل	مختارة	مساعدة
مسائل	مشكلة	مضامين	معالم	مقترح	مواقف
مؤشرات	نبذة	نشاطات	نظريات	نظرات	واضح
واقع	وظيفتها	يعمل	يكون	مقدر	ممارسة

الفصل الرابع

"أل" التعريف والاسترجاع

- مشكلة (أل) الأصلية
- الحروف المتصلة
- واو العطف



الفصل الرابع
"أل" التعريف والاسترجاع

"أل" التعريفية من أعقد القضايا التي حيرت بعض المعجميين العرب في التعامل معها ومعالجتها عند إعداد المعاجم وترتيب الفهارس . إلا أن المكتبيين قد قرروا منذ أمد إغفال "أل" التعريفية في الترتيب الألفبائي خصوصًا إذا وقعت في أول كلمة في الفهارس والكشافات التقليدية .

ومع بداية استخدام الحاسوب في المكتبات العربية برزت "أل" التعريف بمضامين جديدة محيرة وصعوبات أكثر تحديًا يمكن إجمالها في النقاط التالية:

١ – "أل" التعريف زائدة من حيث كونها لا تحمل أية دلالات مهمة في نظام استرجاع المعلومات .

٢ – "أل" التعريف مكونة من حرفين هما الألف واللام يلتصقان مباشرة بأوائل الكلمات العربية مما يستحيل معه وضع "أل" التعريفية ضمن قائمة توقف بحيث يستطيع الحاسوب تمييزها بصورة قطعية . فهناك كلمات عربية وأسماء أجنبية تبدأ بألف ولام أصليتين مثل (الله ، التزام ، إلقاء ، الكترونية ، الفرد ، إلياس) وغيرها كثير . والحاسوب لا يستطيع التمييز بين "أل" الأصلية أو التعريفية لأنها جزء لا يتجزأ من كتابة الكلمة العربية.

٣ – يكثر دخول "أل" التعريف على الكلمات العربية الدالة لدرجة أنه يندر أن نجد نصًا عربيًا خاليًا من "أل" سواء كان النص في لغة طبيعية كما

في عناوين المطبوعات أم في لغة اصطناعية كما في المكانز وقوائم رؤوس الموضوعات .

٤ - يؤثر احتساب "أل" التعريف على تحديد مواضع الكلمات في كشافات التباديل الآلية . كما تؤثر "أل" على استرجاع الكلمات المفتاحية على اعتبار أن مصطلح البحث قد يكون معرفًا بـ "أل" أو غير معرف بها .
فاحتساب "أل" في الترتيب يؤدي إلى تشتت المداخل المتطابقة من حيث احتمال توزيعها في الكشافات تحت حرف الألف ثم حرف اللام وتوزيع المداخل نفسها أو الكلمات الخالية من "أل" على بقية الألفبائية العربية ، سواء كانت المداخل مفاهيم موضوعية أو أسماء مثل :

الصالح ، محمد أحمد	يرتب تحت حرف الالف
صالح ، محمد أحمد	يرتب تحت حرف الصاد
التنمية الإدارية	يرتب تحت حرف الألف
تنمية الإدارة	يرتب تحت حرف التاء

٥ - يؤدي احتساب "أل" في الترتيب إلى تكدس أغلب الكلمات في موضع واحد من الكشاف مما يقلل من فاعلية نظام الاسترجاع الآلي ، ويجعل عملية معاينة المداخل على شاشة المطراف عملية مضنية من حيث أن الباحث سيضطر إلى تتبع الحرف الثالث من بدايات الكلمات المتعاقبة التي قد تستوعب شاشات كثيرة .

- يؤدي احتساب "أل" التعريف إلى ضرورة إدخالها دائما مع كل عملية بحث عند الرغبة في الحصول على مضاهاة موفقة مما يضيف أعباء في البحث لا داعي لها ، أما إذا كانت "أل" مهملة في الترتيب فان

الباحث يبدأ بطباعة مصطلح البحث بـ "أل" أو بدونها وسيحصل على جميع الكلمات المتطابقة لبحثه سواء كانت معرفة بـ "أل" او بدونها كما هي في قاعدة المعلومات .

٧ - يؤثر احتساب "أل" على أداء وفاعلية نظام الاسترجاع العربي من حيث أن الباحث يفترض فيه أن يتوقع بأن مصطلح البحث قد يرد على حالتين : إحداهما معرفة بـ "أل" والأخرى خالية منها مما يضطره إلى عمل حركتين في البحث عن مصطلح ومفهوم واحد . أما إذا استخدم الباحث نظام الاسترجاع وهو لا يعرف أو نسي قضية "أل" التعريف المحسوبة فإن المعلومات المسترجعة ستكون ناقصة .

وأمام هذه المشكلات المصاحبة لاستخدام "أل" التعريف ونظراً لكثرة ورودها في العناوين وفي بدايات الكلمات الدالة فيجب أن لا تكون معالجة "أل" مقيدة بشروط طباعية عند استخدام مفاتيح المطراف أثناء الفهرسة والبحث لأن ذلك يعوق عملية إدخال البيانات ويرهق القائم بالفهرسة ويجعل عملية إدخال البيانات عرضة لكثير من الأخطاء الطباعية . ويمكن معالجة "أل" التعريف بطريقة منطقية وآلية لتحقيق الأهداف التالية :

١ - إهمال "أل" التعريف آليًا .

٢ - ترتيب الكلمات المفتاحية في موضع واحد سواء كانت بـ"أل" أو بدونها.

٣ - استرجاع الكلمات البحثية سواء كانت بـ "أل" أو بدونها .

٤ - عدم التأثير على الشكل الكتابي للكلمات المبدوءة بـ"أل" وبقاؤها متصلة بها.

٥ - التمييز بين "أل" التعريف والألف واللام الأصليتين .

٦ - المعالجة الفردية للكلمات القليلة الملتبسة .

مشكلة "أل" الأصلية :

تبدأ بعض الكلمات العربية والمعربة بألف ثم لام أصليتين مثل (الذي ، التزام ، ألغام ، ألمانيا ، اليوت) . وبغض النظر عن التفسير المنطقي والصرفي فإن "أل" الأصلية ليست من "أل" التعريف في شيء كما توهم بعض المكتبيين العرب ، ذلك أنه لا يمكن أن نبتر "أل" الأصلية فيستقيم المعنى المعجمي أو الدلالي للكلمة المبدوءة بها . أما "أل" التعريف فهي أداة زائدة تلحق الكلمة وعند حذفها في الترتيب لا يتغير مضمون الكلمة بما يحمله من مفاهيم دالة في الاسترجاع مثل : (الإدارة = إدارة ، السعودية = سعودية) . كما يمكن أن تدخل " أل" التعريفية على المبدوء بألف ولام أصليتين إذا كان المفهوم نكرة . وهنا يتغير رسم الحرف المبدوءة به الكلمة إذ تصبح "أل" التعريف + "إأل" الأصلية = (ألف لام ألف لام) مثل : الالتزام ، الإلهام .

وعند دخول "أل" التعريف على المبدوء بألف تصبح بداية الكلمة (أل + أ) = الا ... مثل الأنباء ، الإعلام ونحو ذلك ، وشكل (الا) في بداية الكلمة في الحالتين يعامل معاملة "أل" التعريف من حيث الاستبعاد والترتيب بحيث يكون الترتيب تحت أول حرف أصلي في الكلمة ، إذ يبدأ الترتيب حسب موقع الحرف الثالث .

وما دام أن "أل" التعريف مهملة في الترتيب ونظراً لالتباسها بـ "ال" الأصلية فإن ذلك يستلزم ضرورة التمييز بينهما وتصميم نظام منطقي للمعالجة أو برنامج يهمل "أل" التعريف في البدايات مع إيجاد قائمة استثناء تحتوي جميع الكلمات المبدوءة بألف ولام أصليتين ليقوم الحاسوب ذاتيا باحتسابها عند تبويب الكشاف وعند الاسترجاع .

أما اختيار قائمة الاستثناء للكلمات المجازة Go List الخاصة بـ "أل" الأصلية فيتم على اعتبارات التخصص الموضوعي لنظام الاسترجاع واحتمال ورود الكلمات في سياق العناوين ، وهناك كلمات معجمية كثيرة تبدأ بحرف ألف ولام أصليتين . ومهما استرشدنا بالمعاجم اللغوية فإن حصر الكلمات المبدوءة بـ "أل" أصلية يجب أن يقوم على افتراض أن الكلمات ذات مفاهيم

إله	إلهية	ألسنية	ألفي	إلمام	ألمعية
الله	البان	ألسنيات	ألعاب	إلهام	ألمعي
آلة	التماس	التقاط	إلغاء	ألوان	ألمانيا
ألبرب	الألباني	البيري	الألوسي	ألبانيا	ألمنيوم
آلات	التباس	ألفاظ	ألغاز	الكترون	اليوت
آلي	التزام	إلقاء	ألغام	الكترونية	إلياس
آلية	إلزام	إلقائية	ألمح	ألكتروني	ألياف
إلهيات	التفات	ألقاب	إلماح	الكترونيات	الينوي

شائعة الاستخدام في الفهارس العربية ، أما المعاجم فهي تساعد في الحصر الأولي ومعرفة التصاريف المحتملة للمفاهيم الشائعة ، كما أن هناك كلمات يكثر ورودها بـ "أل" أصلية في العناوين لكنها كلمات فارغة وغير دالة ، ولذلك فهي توضع ضمن قائمة التوقف العامة للكلمات غير الدالة .

وفيما يلي نورد مسردا لبعض الكلمات المبدوءة بألف ولام أصليتين مما قد يرد في عناوين الكتب العربية أو في اللغات الاصطناعية :

واخيراً يجب أن نشير إلى الكلمات المشتركة في الرسم والمختلفة في الدلالة مما قد يشير اللبس عند المعالجة الآلية في النظم التي لا تميز بين حركات

الكلمات، أو بين الألف المهموزة وغير المهموزة عند الترتيب مثل (ألسنية) و (السنية) وهذا يعني أن نستثني بعض الكلمات الملتبسة عند الفهرسة حتى لا يتم بتر أولها كما في الكلمات المزدوجة التالية :

إلهام	فعلها ألهم تبدأ بألف ولام أصليتين
الهام	من مهم تبدأ بحرف الهاء
ألبان	جمع لبن تبدأ بحرف ألف
البان	نوع من الشجر تبدأ بالباء
ألفي	مثنى ألف أو نسبة إليه تبدأ بحرف ألف
الفيء	بمعنى الخراج تبدأ بحرف الفاء
ألغاز	جمع لغز تبدأ بحرف ألف
الغاز	نوع من الوقود تبدأ بحرف الغين .

ونحو ذلك من الكلمات المشتركة في الرسم والمختلفة في الدلالات وفي بدايات الترتيب ، مما يستلزم معالجة خاصة وانتباهاً أثناء الفهرسة الآلية . فإذا جاء بحث بعنوان (استهلاك الغاز في البلاد العربية) فإن الحاسوب لا يفرق بين (غاز) المعرفة بـ "أل" وبين "ألغاز" كما في عنوان (ألغاز وأحاجي) لأن الكلمتين متطابقتان في الرسم ، والمبدوءة بـ"أل" أصلية مختزنة ضمن قائمة التوقف الخاصة بـ "أل" الأصلية . ولذا سيرتب الحاسوب كلمة (ألغاز) بمعنى وقود تحت "أل" التعريف ، إلا إذا تدخل المفهرس ، وميز بين الكلمات المتجانسة ، سواء كانت "أل" أصلية ومحسوبة ، أو كانت تعريفية مهملة في

الترتيب . كما أن عدم وجود قائمة استثناء أو توقف للكلمات المبدوءة بـ "أل" أصلية ستجعل الحاسوب يبتر آلياً "أل" الأصلية من أوائل الكلمات فتصبح (ألغاز) في (ألغاز وأحاجي) بمعنى (غاز = وقود) .

وهكذا نرى أن معالجة "أل" التعريف منطقيًا وآليًا عملية دقيقة ومهمة جدا في الاسترجاع . ومع ذلك نجد أن أغلب برامج المكتبات الجاهزة والمشهورة باللغة العربية لم تعالج قضية "أل" التعريف معالجة شمولية ومقننة .

الحروف المتصلة :

ظلت الحروف المتصلة بأوائل الكلمات المفتاحية تمثل مشكلة في ترتيب الفهارس والكشافات اليدوية شأنها في ذلك شأن "أل" التعريف . وبعض الذين احتسبوا الحروف المتصلة في الترتيب وقعوا تحت تأثير المنطق اللغوي . ومع ذلك فإن بعض المفهرسين والمكشفين الأوائل غلبوا الناحية الاسترجاعية وأهملوا جميع الحروف المتصلة بأوائل الكلمات عند تبويب كشافات القرآن الكريم والحديث النبوي . (الصوينع).

والحروف المتصلة هي حروف الجر التي تشمل اللام والباء والكاف (ل ، ب ، ك) التي تلتصق بمفردات أخرى سواء كانت المفردات دالة مثل (للأدب ، بالتعليم ، كمعرفة) أو كانت الحروف متصلة بمفردات أخرى غير دالة مثل (بها ، لهذا ، كما) .

وقد بين المسح التجريبي لكلمات التوقف أن بعض الحروف المتصلة تتردد بصورة عالية في العناوين وخاصة حرفي (اللام والباء) أما (الكاف والفاء) فهي نادرة الوقوع في العناوين ومن أمثلة الحروف المتصلة الشائعة ما يلي :

ب الإدارة بالمشاركة

ل التدريب للتنمية

ك التدريب كأساس لتنمية القوى العاملة

وتشبه الحروف المتصلة "أل" التعريفية من حيث إنها تتصل ببدايات الكلمات وتكون جزءاً منها ، غير أن الحروف المتصلة متغيرة من حيث إنه يحتمل التصاقها بكلمات تكون أوائلها جميع حروف الهجائية العربية . ومن هنا فإن الحروف الزائدة الملتصقة بأوائل الكلمات تعد كتابياً جزءاً من الكلمات ، ومن ثم يستحيل التمييز الآلي بين الكلمات المبدوءة بحروف أصلية من الكلمات المبدوءة بحروف جر متصلة وزائدة على الكلمة . فالحاسوب لا يستطيع التمييز بين الباء في (بيت ، بلاغة ، بترول) وبين الباء في (بإدارة ، بجامعة) وهكذا بقية الحروف المتصلة الأخرى .

وتؤثر الحروف المتصلة على ترتيب الكلمات الدالة وعلى الاسترجاع من ناحية أن الباحث في كشاف التباديل لا يعرف بأن مصطلحات البحث مجردة أو متصلة بحروف جر . وهذا الوضع يؤدي إلى مشكلات الاسترجاع وتشتت المداخل التي سبقت الإشارة إليها في الكلام عن "أل" التعريف . وقد وقع في هذا اللبس مكنز (الجامعة : مكنز ثلاثي اللغات ...) الذي أصدرته جامعة الدول العربية في تونس عام ١٩٨٧م . إذ نجد أن مكنز (الجامعة) قد رتب الواصفات المقترنة بالباء واللام تحت هذين الحرفين ، ولذا اعتبر هذان الحرفان في الترتيب كأنهما حرفان أصليان من الكلمة ، فكلمة (للأقمار) وجدت مرتبة تحت حرف (اللام) وفي رأس موضوع (أجهزة مضادة للأقمار الصناعية) عند

تباديل الكشاف في المكنز ، ونتيجة لذلك نجد أن كشاف الكلمات المفتاحية خارج السياق يرتب كلمة (بمحرك) في حرف (الباء) وكلمات (للوطن ، للمجتمع، للأسعار ، للمواطنين) قد رتبت تحت حرف (اللام) الجارة . هذا الي جانب أن مكنز (الجامعة) قد ترك بعض الكلمات غير الدالة لتظهر في التباديل مثل (عن ، ما) في الواصفات مثل (الإضراب عن دفع الإيجارات ، التنقيب عن الآثار ، تعليم ما قبل المدرسة) .

وعلى الرغم من صعوبة قضية الحروف المتصلة ، إلا أن معالجتها تكمن في الوصول إلى أنماط منتظمة يترتب عليها تأسيس قاعدة منطقية لاستبعاد الحروف آليا دون تأثير على كلمات أخرى مبدوءة بحروف أصلية مشابهة ، وعند تحليل أوضاع الحروف المتصلة بأوائل الكلمات نجدها تتصل بأوائل الكلمات في ثلاث حالات يمكن على ضوئها معالجة الحروف الملتصقة على النحو التالي :

١ - الاتصال بالكلمات غير الدالة :

إذا اتصلت الحروف بأوائل الكلمات غير الدلالية (بما ، لما ، كما ، به ، بها لبعض ، لأساس ، كأساس) ونحو ذلك فإنه من المنطقي أن نعتبر هذه المفردات كلمات توقف بزوائدها من الحروف الملتصقة . ومن ثم فإن الحاسوب يتولى استبعادها خلال الترتيب والاسترجاع ، ولذا فمن المفترض أن تؤخذ هذه المفردات في الحسبان عند تجميع كلمات التوقف ، وإلا فإن الحاسوب سوف يعتبرها في التباديل لأنه لا يفرق بين الكلمات إذا اختلفت تصاريفها . فالحاسوب لا يميز بين (هذا) و (لهذا) أو بين (آراء) و (لآراء) ونحو ذلك .

٢ - اتصال الحروف بالمعرف بـ « أل » :

إذا اتصلت الحروف بأوائل الكلمات المبدوءة بـ " أل " التعريف أصبحت هناك نمطية منتظمة في تعاقب الحروف كما يلي : (للإدارة ، للتعليم ، بالمعرفة ، بالأدب) فهذه الكلمات مبدوءة بحروف (ل ل ، ب ا ل) . وهنا يمكن القول إنه إذا وردت الحروف المتصلة مع "أل" التعريف في أوائل الكلمات الدالة فتحذف الحروف المتصلة مع "أل" التعريف وسيتم ترتيب الكلمات حسب أوائلها الأصلية .

لكن هذه القاعدة قد تثير اللبس في الكلمات المبدوءة بحروف أصلية مطابقة لنمط تتابع حروف الجر مع "أل" التعريفية مثل :(اللغة ، للجنة ، للاجئين ، بالغ، بالغون ، بالغات ، بالية ، باليه ، بالون) ونحوها . فهذه الكلمات مبدوءة بحروف (ل ل ، ب ا ل) فلا يمكن حذف اللام الثانية في كلمة اللجنة آلياً ، كما لا يمكن حذف الحروف الثلاثة من (بالغ) وما شابهها لأنها مبدوءة بحروف أصلية، وما يقال عن هذه ينطبق على (الكاف) أو حتى (فاء) الجر على الرغم من ندرة وقوع الكلمات الملتبسة في العناوين ، مثل :(فالوب ، فالحون ، فالح) والكاف مثل : (كالح ، كالحين ، كالحات ، كالدونيا) . فهذه الكلمات مبدوءة بحروف أصلية لكنها تشبه أغلب الكلمات المبدوءة بحروف الجر (الفاء ، الكاف + "أل" التعريف) .

والحل المناسب لهذه المعضلة أنه يمكن استثناء الكلمات المبدوءة بحروف أصلية وحصرها ضمن قوائم التوقف المجازة أو المستثناة Go List ليتمكن الحاسوب من استثنائها آلياً خلال المضاهاة والاستبعاد .

٢ - اتصال حروف الجر بكلمات دالة غير معرفة بـ "أل" :

ويشمل ذلك جميع الكلمات العربية الدالة التي تكون مبدوءة بأي حرف من حروف الجر ، شرط أن تكون غير معرفة بـ"أل" مثل (بإدارة ، بتعليم ، لمعرفة ، كتدريب) ونحو ذلك . وهذه الحالة لا يمكن علاجها آلياً ، وإنما تتطلب المعالجة الفردية واليدوية باستخدام رموز موحدة تمكن الحاسوب من إلغاء قيمة الحرف الأول في الكلمات المفتاحية ، مع بقاء رسم الكلمات بحروفها المتصلة ، وبذلك فلن تؤثر الحروف المتصلة على الترتيب والتباديل أو تشتيت المداخل مما ينعكس سلبًا على الاسترجاع .

واو العطف :

واو العطف حرف منفصل في الكتابة العربية لكنه يتردد بصورة عالية جداً في العناوين وفي النصوص الأخرى . وتؤثر واو العطف على ترتيب الكلمات وفرزها في كشافات التباديل إلى حد يستحيل معه استرجاع الكلمات الدالة المعطوفة إذا عرفنا أن واو العطف تتصل بما بعدها في الطباعة بحيث لا يكون هناك فراغ بين الواو وما بعدها حيث يمكن للحاسوب تمييز واو العطف عن الكلمة المعطوفة .

وعلى الرغم من أن الإنسان يميز دلاليًا واو العطف عن الكلمة المبدوءة بواو أصلية مثل (وراثة ، وثائق) إلا أن الحاسوب بدون وجود فراغ ، يعتبر واو العطف وما بعدها كلمة واحدة . فالبحث عن كلمة (التعليم) في عنوان (التربية والتعليم) لن يؤدي إلى استرجاع مصطلح البحث ، إلا إذا كان مكتوبا عند الاسترجاع بنفس طريقة الإدخال في الفهرسة هكذا (والتعليم) بدون فراغ . لكن الحاسوب سيسترجع نفس مصطلح البحث لو كان في عنوان آخر (التعليم

والتربية) وهذا الأسلوب يضطرنا إلى الاحتياط وإدخال مصطلح البحث على احتمالين ، مرة مجردا ومرة مع واو العطف عند الرغبة في استرجاع معلومات كاملة . ولكن ذلك لا يعد مرونة في نظام الاسترجاع ، ولذا فإن الحل الأمثل أن نجعل واو العطف كلمة توقف ثم نفصلها عما بعدها بفراغ دائما عند الإدخال، وهذا يمكن الحاسب من تمييزها عند المضاهاة في تباديل الكلمات المفتاحية، ومن ثم تكون القاعدة : أنه إذا وردت واو عطف يسبقها فراغ ويلحقها فراغ فإنها تهمل في الترتيب والاسترجاع ، وإذا كان شكل النص المطبوع غير مقبول لانفراد الواو بين فراغين فيمكن حل هذه المشكلة في برنامج الطباعة الذي يتولى لزّ واو العطف بما بعدها على المستخرجات المطبوعة .

كما يمكن معالجة واو العطف بطريقة منطقية عامة عند الرغبة في عدم فصلها عما بعدها أثناء الطباعة ، وذلك بشرط أن تكون متصلة بمعرف بـ "أل" مثل (والسياسة ، والقانون ، والاقتصاد ، والهندسة) . ومع ذلك يجب إعداد قائمة استثناء للكلمات التي قد تبدأ بحروف (وأل) أصلية مثل (والله ، والد، والدين ، والي) ونحو ذلك ، كما لا بد من ملاحظة التباس كلمتي (والدين ، والدّين) إحداهما مثنى أو جمع (والد) المبدوءة بواو أصلية وكلمة (الدّين) بتشديد الدال هي كلمة معرفة بـ "أل" . فاللبس يظهر إذا كانت الكلمة الأخيرة معطوفة بواو . ومع ذلك فان هذه القاعدة المنطقية لمعالجة واو العطف المتصلة بالمعرف بـ (أل) لن تغني عن المعالجة الخاصة للمعطوفات غير المتصلة بـ "أل" التعريف التي تستلزم فصل الواو بفراغ عند الإدخال . لكن هذه القاعدة تخفف معالجة واو العطف عند نسيان الفراغ الفاصل بين الواو وبين الكلمات المعطوفة إذا عرفنا أن نسبة المعطوفات المعرفة بـ "أل" عالية جداً في العناوين وفي اللغات الاصطناعية .

الفصل الخامس

الترتيب والمواصفات

- الترتيب الألفبائي
- المحارف العربية
- التحليل الصرفي
- بتر الكلمات



الفصل الخامس
الترتيب والمواصفات

الترتيب الألفبائي :

قواعد التصفيف الألفبائي متفاوتة بين المكتبات العربية وفي الفهارس والكشافات المطبوعة ، ولم تحسم مسألة الترتيب حتى بعد استخدام الحاسوب في أعمال المكتبات خصوصا ما يتعلق بالألف والهمزات وعلامات التشكيل وكلمات التوقف غير الدالة التي جاءت بمشكلات جديدة تحتاج إلى مزيد من الدراسات والتأصيل بما يوافق احتياجات البحث والاسترجاع في المكتبات ومراكز المعلومات .

ويجب أن تعتمد نظم المكتبات على الترتيب الألفبائي الصارم الذي أساسه الكلمة حيث تقف توافقات المضاهاة الحرفية بنهاية الكلمة الأقصر مقارنة بينها وبين الكلمة التي تسبقها والتي تليها ، أي أن الفراغات تسبق في الترتيب وهي التي تحدد بداية ونهاية الكلمة مثل :

أ . نديم

أب

آبار

إبر

أبراج المدينة

إبراهام

إبراهيم أحمد

ابراهيم يحيى

إبريق

آثار

أويك

وتؤثر قواعد ومواصفات الترتيب المتبعة على ترتيب المداخل وعرض الكشافات وكذلك على دقة الاسترجاع خصوصًا ما يتعلق بالهمزات ، غير أن المكتبات تتجه على مساواة أشكال الألف المهموزة والممدودة والمفردة وإهمال حركات التشكيل كما فى القواعد التالية المتبعة في قائمة رؤوس الموضوعات العربية لمكتبة معهد الإدارة العامة :

* تعامل الألف الممدودة معاملة الألف العادية مثل :

أثاث المكاتب

آثار الالتزام

* تعد الهمزة على الألف ألفا مثل :

التأديب

التاريخ

التأمين

* وتعد الهمزة المفردة على السطر ألفا مثل :

القراءة

القرآن

* وتعامل الهمزات على الألف والواو والياء بنفس ترتيب الحروف التي تحملها (أ ، و ، ي) حسب مواقعها المعتادة من الترتيب سواء كانت مهموزة أو مجردة (ء أ إ آ ا = ا ، ؤ = و ، ئـ ، ي) .

* كما تعامل التاء المربوطة مفردة أو متصلة معاملة الهاء في الترتيب مثل:

مراكز المعلومات

المرأة

* كما تسبق الياء غير المنقوطة الياء المنقوطة في نهاية الكلمة مثل :

على ، علي ، روى ، روي

وفي النظم الآلية يجب ترتيب الكلمات وفقا لتسلسل الألفبائية العربية بشكل صارم ، وعدم احتساب علامات الترقيم مثل النقطة والفاصلة والشرطة في الترتيب وعدم إحلال الحروف المستخدمة محل علامات الترقيم أو عمل استثناءات منطقية كما في الترتيب اليدوي ، مثل الترتيب الزمني الذي يتعارض مع الألفبائية كما في الأمثلة التالية :

الشعر العربي - صدر الاسلام

الشعر العربي - العصر الأموي

الشعر العربي - العصر الجاهلي

المؤتمر العربي الاول

المؤتمر العربي التاسع

المؤتمر العربي الثالث

المؤتمر العربي الثامن

المؤتمر العربي الثاني

المؤتمر العربي الخامس

المؤتمر العربي الرابع

المحارف العربية :

من الإشكالات التي تربك الترتيب الألفبائي وتعيق المضاهاة الدقيقة أثناء الاسترجاع مسألة التقيد بالمواصفات المتبعة في تمثيل المحارف العربية في الحاسوب . ومن المواصفات العربية المستخدمة (الشفرة العربية الموحدة) أو غيرها . وتستلزم المواصفات القياسية تمييز أشكال الألف في الترتيب والاسترجاع حسب اختلاف مواضع الهمزة في أول الكلمة سواء كانت مفتوحة أو مكسورة أو مضمومة أو ممدودة . وكذلك التمييز بين الهاء في آخر الكلمة مع التاء المربوطة .

ورغم الأهمية اللغوية الأكيدة للتمييز بين أشكال الهمزات والهاء والتاء وحركات التشكيل على حروف الكلمات في نظم استرجاع النصوص وفي الدراسات اللسانية والأدبية وما يتفرع عنها من دراسات نصوصية إلا أن لتمييز أشكال الألف في بداية الكلمة تأثيرات سلبية على نظم استرجاع المعلومات في المكتبات وفي التطبيقات التوثيقية والإدارية ، ويمكن تلخيص أبرز المشكلات الناجمة عن التقيد بمواصفات المحارف العربية القياسية فيما يلي:

١ - إن نظم المعلومات في المكتبات ومراكز المعلومات لا تتعامل في التخزين والاسترجاع مع نصوص ثابتة وإنما تتعامل مع بيانات ومعلومات نشطة يتم تحديثها واسترجاعها باستمرار .

٢- صعوبة التقيد الصارم بالشروط الإملائية الدقيقة أثناء إدخال البيانات بكميات ضخمة خصوصا وأن معظم المفردات لا تثير اللبس ويمكن تمييز دلالاتها مع وجود الأخطاء الإملائية فيها . بينما نجد أن الكلمات التي تثير اللبس أو تؤدي إلى الأخطاء الإملائية أو الدلالية يمكن السيطرة عليها بسهولة من خلال الالتزام بالقواعد الإملائية الصحيحة

وليس من خلال التقيد بالشروط التي تفرضها مواصفات المحارف المتبعة في الحاسوب .

٣ - ضعف المهارات الإملائية لدى معظم الموثقين وعلى الأخص مدخلي البيانات كما أن التشدد في تدريبهم وإلزامهم بالمواصفات الدقيقة عملية مكلفة وتؤدى إلى خفض الإنتاجية .

٤ - ضعف المهارات اللغوية أو تفاوتها بين مستخدمي نظم المعلومات من القراء والباحثين الذين تختلف قدراتهم وتخصصاتهم واحتياجاتهم المعلوماتية . هذا علاوة على استحالة إلزام جمهور المستفيدين بالتقيد بالمواصفات الإملائية مما يعكس آثارا سلبية على استخدام نظام المعلومات في البيئة المفتوحة مثل المكتبات .

وبمقارنة الميزات مع عيوب التقيد بالمواصفات المستخدمة لتمثيل المحارف في الهمزات نجد أن الحل المناسب هو أن يكون نظام الاسترجاع مرناً في تقبل الأشكال المختلفة للإملاء سواء كان ذلك نتيجة خطأ أو سهو أو تفاوت الإملاء بين الإدخال والبحث لا سيما بأن الهمزات والتاء المربوطة أو الهاء المشابهة لها من أكثر الأخطاء الإملائية تكرارا ويضاعف ذلك احتمال وقوع الخطأ في حالتين:

أ - وجود خطأ أو تفاوت في إدخال البيانات أثناء الفهرسة والتكشيف .

ب - وجود خطأ أو تفاوت في الإدخال أثناء البحث والاسترجاع .

فإذا كانت أشكال الألف الابتدائية (ا = أ = إ) متساوية القيمة في الترتيب والاسترجاع فإن العبء يقع على الحاسوب في تمييز أشكال الألف أو التاء المربوطة عند الاسترجاع والترتيب سواء كانت مدخلة بالشكل الذي استخدمه الباحث أم لا . فالحاسوب سوف يسترجع إدارة ، ادارة ، الاردن ، الأردن ، صحة ، صحه) سواء كانت الكلمة مهموزة أم لا وسواء كانت منقوطة أم لا .

ومع ذلك يجب العناية في دقة إدخال الكلمات القليلة التي تثير اللبس سواء كانت مبدوءة بالألف المهموزة كما ذكر ذلك في مشكلة "أل" الأصلية أو كانت الكلمات مختومة بالتاء المربوطة .

ورغم أن حالات الالتباس محدودة في مجال المكتبات والمعلومات إلا أنه يوجد بعض المفاهيم الاسمية المختومة بالتاء المربوطة تتشابه مع الفعل الماضي المتصل بالهاء أو الاسم المضاف إليها لو حدث خطأ إملائي في إغفال النقطتين على التاء أو جرى تحريف موقع الهمزات مثل :

آخرة (ما بعد الموت)

آخره (نهايته)

إدارة (إدارة عامة)

أداره (فعل ماضٍ + مفعول به)

أسرة (عائلة)

أسره (من الأسر)

إقامة (إقامة الصلاة ، رخصة الأجانب)

أقامه (فعل ماضٍ + مفعول به)

جدة (حداثة ، اسم مدينة ، أم الأب والأم)

جده (أب الأب + ضمير الغائب المضاف إليه)

زكاة (أحد أركان الإسلام)

زكاه (من التزكية فعل + مفعول به)

سيرة (تاريخ)

سيره (من السير)

طلبة (تلاميذ)

طلبه (فعل ماضٍ + مفعول به مضاف ومضاف إليه)

معاهدة (اتفاقية)

معاهده (معاهد + هـ)

والخلاصة : أنه مهما كانت مرونة نظام المعلومات في تقبل البحث والاسترجاع أو الترتيب بالأشكال الإملائية المتفاوتة إلا أنه يجب أن يكون هناك حد أدنى لضبط جودة المعلومات والتقيد بالقواعد اللغوية الأساسية أثناء الفهرسة وإدخال البيانات خصوصاً في المسائل الإملائية التي تثير اللبس أو تؤثر على صحة المعلومات . فالدعوة لمساواة أشكال الألف المهموزة في تمثيل المحارف وفي الاسترجاع ليست موجهة نحو القواعد الإملائية المتبعة في الكتابة العربية ولا لكافة نظم المعلومات وإنما هي موجهة لنظام المعلومات المتبع لمعالجة واسترجاع المعلومات في المكتبات ومراكز المعلومات على الأقل في الوقت الراهن حتى يتم تطوير برامج قوية لتشكيل الكلمات العربية تلقائياً في سياقاتها حسب دلالاتها وتركيبها اللغوي .

التحليل الصرفي :

يعتمد برنامج المحلل الصرفي الذي ابتكرته الشركة العالمية للإلكترونيات في برنامج القرآن الكريم وبرنامج صحيح البخاري على المبدأ اللغوي وهو أن أساس بنية الكلمة العربية ثلاثة أحرف تمثل دائما بحروف (ف ع ل) باعتبارها الجذر الذي يطرأ عليه متغيرات صرفية أو زوائد صرفية أمامية وهي السوابق وخلفية وهي اللواحق ، جميعها يطلق عليها اللواصق مثل (مكتوب) و (كتابهم).

ويتم استرجاع النصوص بالاعتماد على المحلل الصرفي الذي يتولى ربط المفردات المختلفة الصيغ بجذورها مثل (كتب ، الكتب ، يكتبون ، كاتبون ،

اكتبن ...) التي يمكن استرجاعها تحت الجذر (كتب) بالإضافة إلى إمكانية استخدام الكلمات المختلفة حسب صياغاتها المتفاوتة مع ما يتصل بها من سوابق مثل حروف الجر والعطف وسين الاستقبال أو لواحق مثل ضمائر الرفع والنصب والجر أو ياء النسب . كما أن دقة الاسترجاع تعتمد على التقيد بحركات التشكيل (الضمة ، الفتحة ، الكسرة والشدة) ، وينعكس غياب أو وجود التشكيل على دقة الاسترجاع بما يحدثه من لبس في تحديد أصل المفردات وتعدد جذورها المحتملة فمثلا (أنزل) لها سبعة احتمالات في تشكيل الحركات ، كما أن كلمة (قل) قد يكون جذرها المضعف (قلل) وعندها تكون الصيغة الصحيحة للكلمة هي (قلّ) أو قد تكون الكلمة ناتجة عن الجذر (قول) وعندها تكون الكلمة الصحيحه (قُلّ) . ولذا يستلزم تعيين الجذر المطلوب بدقة وضبط حركات التشكيل بشكل صارم من أجل الحصول على مضاهاة دقيقة وإلاّ ستكون نتائج البحث تقريبية .

ورغم صعوبة التمسك بشروط التحليل الصرفي عند الاسترجاع إلاّ أن ذلك ضروري للحصول على مضاهاة دقيقة خصوصاً وأن البحث يتم في نصوص محددة وغير قابلة للإضافات أو التحديث من قبل المستفيد النهائي هذا علاوة على أن للمفردات المستخدمة في البحث أهمية ودلالات لغوية تتجاوز الدلالات الاصطلاحية أو الموضوعية . ذلك أن دلالات المفردات متغيرة حسب سياقاتها وحركاتها في اللغة الطبيعية سواء كان البحث في النص القرآني أم في الحديث النبوي . ولذا نجد أن البحث يتم بالمفردات على اختلاف صيغها المتولدة من الجذور سواء كانت الكلمات أسماء أو أفعالاً وغيرها مع وضع حركات التشكيل المناسبة على الحروف حسب المستويات البحثية التالية :

١- مستوى الكلمة :

وهذا يعني أن الباحث يرغب في استرجاع نصوص الآيات والأحاديث التي تحوي كلمة مطابقة تماماً للكلمة التي استخدمها الباحث .

٢- الكلمة مع اللواصق :

ويعني ذلك إمكانية استرجاع الكلمة المطلوبة مهما وردت مرتبطة بالسوابق واللواحق سواء كانت متصلة بأل التعريف أو همزة الاستفهام أو الضمائر أو الفاء والسين أو حروف العطف ، فالبحث بكلمة (آمن) يؤدي إلى استرجاع النصوص التي وردت فيها كلمات (آمنون ، آمنات ، الآمن ، فالآمنون ، والآمنات ، آمنوا ، الآمنين) .

٣- مستوى الجذر :

على هذا المستوى يستطيع الباحث استرجاع النصوص التي وردت فيها اشتقاقات الكلمة المحددة المستخدمة في البحث . فإذا أدخلت كلمة (مؤمن) يسترجع برنامج المحلل الصرفي جميع الكلمات التي فيها اشتقاقات الجذر(آمن) مثل (مؤمن ، مؤمنين ، يؤمنون ، وآمنين ..... إلخ) فهذا النمط من البحث يجمع الاشتقاقات حسب الكلمة البحثية وليس حسب الجذر .

٤- البحث بالجذر فقط :

بالبحث على مستوى الجذر فقط يستطيع الباحث استعراض النصوص التي تتضمن اشتقاقات الجذر المستخدم في البحث . ولا يقبل هذا النمط من البحث إلا الجذور . فإذا أدخلت الجذر (رحم) فسيتم استرجاع الآيات أو الأحاديث التي تتضمن كلمات مثل (الرحمن ، الرحمة ، يرحم ، رحيم ، يرحمون، الأرحام) .

التحليل الصرفي ونظم المعلومات :

إن استخدام جذر الكلمة في التنظيم والاسترجاع ما هو إلا محاولة للتغلب على صعوبة تفاوت أشكال الكلمات حسب صياغاتها اللغوية والاشتقاقية أو الزوائد التي تطرأ على الكلمات . والهدف من استخدام الجذر هو الرغبة في جمع شمل الكلمات ذات التصاريف المتفاوتة تحت مدخل واحد سواء كان ذلك في معجم أو في كشافات نصوص القرآن الكريم والحديث النبوي وغيرها من نصوص أدبية . ورغم أن هذا الأسلوب قد ينفع في المعاجم وفي التطبيقات اللغوية ودراسة النصوص لمختلف الأغراض إلا أنه لا يصلح للاستخدام في تنظيم المعلومات واسترجاعها لعدة أسباب كما يلي :

١- إن دلالات المفردات الحرة تختلف حسب تفاوت صيغها الصرفية إلى درجة ينسلخ فيها المعنى الاصطلاحي عن المعنى اللغوي الأصلي للكلمات .

٢- إن اللغة المستخدمة لضبط مصطلحات المعلومات مقننة بشكل جيد أو دقيق في معظم حقول المعرفة حتى ولو كانت مكتوبة بلغة طبيعية ونصوص كاملة .

٣- إن أساس التحليل الصرفي يعتمد على التحليل اللغوي المجرد للألفاظ حسب مضامينها ودلالاتها المختلفة في نصوص وسياقات محددة سواء كان التطبيق على نصوص القرآن أو الحديث النبوي أو الشعر أو نحوها.

٤- إن لم شتات الاشتقاقات والتصاريف المختلفة لجذر كلمة لا يعني بالضرورة الوصول إلى دلالات مفهومية أو اصطلاحية مستخدمة في

تنظيم المعلومات مع أن ذلك قد يصدق في التحليل النصي واللغوي للألفاظ المستخدمة في سياقات نص أو نصوص طبيعية محددة .

٥- إن وحدة البحث والاسترجاع ونقل المعلومات بين وسائط ونظم المعلومات في المكتبات ومركز المعلومات وفي البيئة الإدارية هي المصطلحات ذات الدلالات المشتركة بين المرسل والمستقبل وليس الألفاظ والمفردات التي تتغير معانيها حسب سياقات النصوص المختلفة . فمجال التحليل الصرفي موجه نحو الألفاظ والنصوص القابلة لتفسيرات جديدة وليس المعلومات الموضوعية .

٦- ومن الناحية العلمية والسلوكية فإن التحليل الصرفي بما يمتاز به من دقة وصرامة في الضبط والحركات ينفع في التطبيقات والبحوث العلمية حسب طبيعة النصوص المختزنة في نظام المعلومات لكنه غير فعال في بيئة الاسترجاع المفتوحة على جمهور متنوع من المستفيدين كما في المكتبات ومراكز المعلومات.

بتر الكلمات :

البتر وسيلة للتحكم بأجزاء المفردات وتوافقات الحروف في أول الكلمة وفي آخرها أو حتى في وسطها . ويستخدم البتر للبحث في اللغات الطبيعية والنصوص الحرة ، فهو مفيد للبحث بجزء من الكلمة أو للتخلص من اللواصق التي تتعاقب على أصل الكلمة أو شكلها المجرد حسب حالاتها الصرفية والإعرابية المختلفة سواء كان البتر في صدر الكلمة أو في عجزها أو هما معاً بهدف الحصول على أكبر قدر ممكن من المضاهاة المحتملة أثناء الاسترجاع .

ويتم البحث باستخدام جزء مهم من الكلمة البحثية يتم اختياره بعناية حسب تقدير الكلمات الدلالية المحتملة مع وضع علامة البتر الخاصة قبل أو بعد الجذر أو بالأحرى مجموعة الحروف المرغوب اعتمادها كمادة أو أصل لتوليد الكلمات المرغوبة . وعلامة البتر ما هي إلا رمز مبرمج يوضع جنب الكلمة أو جزء من الكلمة ليتمكن الحاسوب من استرجاع كافة الكلمات العربية المختزنة والمحتمل تأليفها بإضافة الحروف الممكنة مكان علامة البتر من أجل تكوين كلمة تامة سواء كان البتر أمام الكلمة أو خلفها مثل :

طريق ×××

طريقة ، طريقتان ، طريقها ، طريقتها ، طريقنا، طرقهم ...

××× طريق

بطريق ، كطريق ، لطريق ...

والبتر الأمامي مفيد للتخلص من السوابق الزائدة مثل " أل " التعريف وحروف الجر والعطف الملتصقة ونحوها إذا كانت الحروف الزائدة ليست مدرجة ضمن قائمة التوقف أو للتخلص من الحروف الأصلية مثل سوابق كلمة (درس) مثل (أدرس ، تدرس ، مدرس ، ندرس ، يدرس) . كما أن البتر الخلفي مفيد مع المفردات التي تولد كلمات كثيرة ومتفاوتة باختلاف نهاياتها سواء كانت ملحقاتها زيادات من الضمائر المتصلة أو ياء النسب أو بسبب الجمع والصفة في مختلف الحالات الإعرابية . على أن البتر في البحث الموضوعي في كشافات العناوين والمستخلصات باللغة العربية لا يعتمد على جذر الكلمات أو أصولها الصرفية ، وإنما يفترض أن يحدد الباحث كلمة كاملة

مجردة أو أجزاء من أوائل الكلمات الدلالية في تصاريفها الكاملة . فالمضاهاة تتم بالربط اللاحق المعتمد على رسم الكلمة بخلاف الكشافات التي تعتمد على الربط المسبق بين جذر الكلمة واشتقاقاتها المختلفة كما هو الحال في كشافات نصوص القرآن والحديث أو في المعاجم اليدوية منها والآلية التي تعتمد على أسلوب التحليل الصرفي . ففي هذه الحالة يتم الربط بين الجذر وما يتفرع عنه من كلمات أثناء الإدخال ويتم الاسترجاع وفق ذلك .

أما في أسلوب البتر فيتم الربط بين توافقات الحروف المستخدمة كمصطلح للبحث مع ما يضاف إلى الكلمة من حروف زائدة أو أصلية تظهر في التركيب اللغوي أثناء المضاهاة والاسترجاع . وعلى هذا الأساس فإن البتر في نظم استرجاع المعلومات يعتمد على تجانس حروف الكلمة وليس على جذرها الصرفي الذي تتغير حروفه حسب تغير حروف العلة الثلاثة (ا ، و ، ي) مع اختلاف التصريف . فالربط المسبق للكلمات بجذرها يتيح استرجاع الكثير من الكلمات المختلفة في تصاريفها مثل (دار ، دوائر ، ديار ، دير) ونحو ذلك تحت مادة (دور) بينما البتر لا يتيح قدر ذلك وإنما يتيح استرجاع كلمات (أدور ، تدور ، مدور ، ندور ، دوري ، دوريات) لوجود حروف متوافقة في هذه الكلمات مع الكلمة (دور) في حالتي البتر الأمامي والخلفي .

ورغم أن أسلوب البتر مناسب للبحوث في اللغات الطبيعية أكثر من اللغات الاصطناعية أو حتى في النصوص المقننة إلا أن استخدامه في اللغة العربية يولد مفردات كثيرة ومتفاوتة من حيث الدلالة والتصريف والإعراب بشكل كبير مما يؤدي إلى وقوع بعض المشكلات الدلالية في الاسترجاع على النحو التالي :

١- استبعاد بعض الكلمات الدالة القريبة من الكلمة البحثية نتيجة اختلاف صيغ المصطلحات المهمة مع ترتيب حروف الكلمة المبتورة .

٢ - ارتفاع نسبة الاستدعاء أثناء الاسترجاع مع وجود نزر يسير من الكلمات الدالة المرغوبة .

٣ - استرجاع كلمات لا علاقة لها البتة بالكلمة البحثية بسبب توافق الحروف في أوائل الكلمات المبتورة من الخلف .

فالبتر الأمامي والخلفي لكلمة (××× علم ×××) بدون نظر للحروف الزائدة مثل حروف الجر والعطف والتعريف ، قد يولد حوالي أربعين كلمة محتملة مع ما يطرأ على الكلمة من حالات الاسم المفرد والمثنى والجمع المذكر والمؤنث في حالات الرفع والنصب والجر ، ومن الأفعال في حالات الماضي والمضارع والأمر مع اتصال الفعل بالضمائر في كل الحالات الإعرابية . فالحروف الأصلية المتغيرة تشمل (أ ، ت ، م ، ي ،) والتي يمكن أن تلتصق بصدر الكلمة ، أو الحروف (ا ، ت ، م ، ن ، هـ ، و ، ي) والتي يمكن أن تلحق عجز الكلمات المتولدة من الأصل أو الكلمة البحثية (علم) سواء جاءت اللواحق مفردة أو ثنائية أو أكثر مثل (اعلم ، اعلمي ، تعلمون ، علماء ، علمي ، علمانية ، علمائهم ، معلم ، معلمين ، معلمات ...) ونحو ذلك . ولكن سوف يتم استبعاد بعض الكلمات الدالة رغم أنها متولدة من الجذر (ع ل م) مثل (عالم ، عالمة ، علوم، معلومة ، معلومات ، معلوماتية) وغيرها وذلك بسبب اختلاف نسق حروف (علم) عن الحروف في الكلمات المستبعدة التي زيدت بحرفي الألف والواو في وسط أو نهاية الكلمة .

وحيث أن البتر على اختلاف مواقعه موجه للتعامل مع المفردات التي تتفاوت معانيها باختلاف صيغها داخل اللغة الطبيعية بغض النظر عن الدلالات الموضوعية ، فإن حدود معاني المفردات المتولدة من البتر متقاربة من الناحية المعجمية رغم زخم المفردات واختلاف الاشتقاقات . ومرد ذلك يعود إلى غنى ومرونة اللغة العربية في مجال التصريف والاشتقاق والزيادات التي تسبق الكلمات وتلحقها نتيجة اختلاف تصاريفها ودلالاتها المعجمية دون أن ينعكس ذلك على توليد مفاهيم اصطلاحية متقاربة أو موضوعية مهمة ضمن سياق الكلمة الواحدة المبتورة .

أما من جهة البحوث اللغوية والنصوصية فإن للبتر اللاحق أو المسبق الذي يعتمد على الجذور أهمية كبيرة في حصر المفردات ودراسة دلالاتها وسياقات استخدامها في النصوص المختلفة . ولذا فإن حركات التشكيل والهمزات مهمة جدا مع استخدام البتر في النصوص الكاملة المشكولة بحيث توضع حركات التشكيل والهمزات في مواقعها المناسبة على الجزء المتبقي من الكلمة المبتورة المستخدمة في البحث .

ولو جرى بتر كلمة (إدار xxx) في حالة البتر الخلفي فسوف تولد كلمات دالة مثل (إدارة ، إدارات ، إداري ، إداريون ، إداريات ، إدارية) مع احتمال استرجاع كلمة أخرى دالة قد لا يكون لها علاقة متوقعة مثل (أدارسة) في حالة إهمال الهمزة في البحث . وكلما قلت الحروف المبقاة من الكلمة المبتورة أو كان البتر مزدوجا تضاعفت نسبة الاستدعاء على حساب التحقيق وزاد الاحتمال في استرجاع مفردات كثيرة مشتتة ، فالبتر المزدوج لكلمة (xxx دار xxx) يؤدي إلى استرجاع مفردات مضاعفة متفاوتة الدلالة بعضها متوقع والآخر غير متوقع مثل (إدارة ، إداري ، إدارات ، دارات ، دارين ، دارس ، مدارس ،

جدار، أقدار ، إهدار ، رادار ، غدار ، مدارج ، مدارات ، مقدار ... وغير ذلك) .

والخلاصة : أن البتر من أساليب البحث الآلية التي يمكن استخدامها للحصول على بعض الكلمات الدالة المختزنة في قاعدة البيانات مع احتمال استرجاع مفردات كثيرة متشابهة في الرسم وإن كانت غير مرغوبة من الناحية الدلالية . كما أن فعالية البتر تعتمد على خبرة الباحث في طبيعة الكلمات المختزنة ودلالاتها المتوقعة وحجمها في قاعدة البيانات . فبعض الكلمات قد تكون مصطلحات معربة أو أسماء جامدة لا ترد إلا في صيغة واحدة أو صيغ محدودة ومتوقعة ، فالبتر في هذه الحالة قد يكون وسيلة سريعة لاسترجاع هذه الكلمات وغيرها من المفردات التي ترد بصيغ موحدة وفي موضع واحد من الترتيب الألفبائي خصوصاً في حالات البتر الخلفي حين تبقى أوائل الكلمات موحدة في مختلف الحالات الإعرابية . كما أن البحث الاستطلاعي بكلمة أو جزء منها في الكشافات قد يغني عن البتر الأمامي ، فالبحث بكلمة (مال) مثلا يتيح استعراض المداخل التي تتضمن الكلمات التالية في الكشاف التباديلي :

مال مالك

مالاوي مالكوم

مالت مالكي

ماله مالي

مالح ماليزيا

مالديف

الفصل السادس

التباديل والاسترجاع

- كشافات التباديل
- عرض المداخل
- الترتيب والاسترجاع
  * مداخل العناوين
  * مداخل الأسماء
  * المستخلصات



الفصل السادس
التباديل والاسترجاع

كشافات التباديل :

تعتمد كشافات الكلمات الدالة في العنوان أو في غيره من النصوص على مبدأ التباديل . والتباديل هو " الترتيب وفق جميع أساليب الترتيب المحتملة " كما ينقل Suihart . وتعود فكرة التباديل إلى مكتبي بريطاني يدعى (أندريا كريستادورو) حيث قدم عام ١٨٥٦م مبدأ تباديل الكلمات المفتاحية في العنوان باستخدام الوسائل التقليدية أو اليدوية . ولكن البداية الحقيقية لهذا النوع من الكشافات قد تمت باستخدام تقنية الحاسوب وذلك في أواخر عام ١٩٥٨م على يد الدكتور (هانز بيترلون) الذي كان يعمل مع شركة (أ بي إم) في واشنطن .Artindi Doyle

ويستخدم التباديل في كشافات العناوين والنصوص المحسبة حيث تطورت نماذجه إلى أشكال متعددة أبرزها كشافات الكلمات الدالة في السياق KWOC وكشاف الكلمات الدالة خارج السياق kwoc وأشكال أخرى مطورة . (الصوينع) .

ويمكن تعريف كشاف الكلمات الدالة في السياق بأنه مسرد ألفبائي للكلمات الدالة أو المفتاحية التي تقع في عناوين الوثائق المكشفة حيث يتم

ترتيب الكلمات الدالة في مواقع متعددة من الكشاف حسب أوائلها من الحروف ، فإذا كان العنوان في الفهرس البطاقي يرتب ألفبائياً حسب أول كلمة في العنوان وفي مكان واحد من الفهرس فإن العنوان نفسه يمكن أن يرتب في مواقع ألفبائية متعددة من الكشاف الآلي حسب عدد الكلمات المفتاحية أو الدالة في نص السياق من العنوان أو غيره . ويمكن إيضاح ذلك في سياق العناوين التالية :

١ - تاريخ الإدارة السعودية .

٢ - ملاحظات على الإدارة في السعودية .

٣ - نبذة موجزة عن تاريخ الإدارة في السعودية .

فنلاحظ في الترتيب الاعتيادي للعناوين الثلاثة أنها وقعت تحت حروف (ت، م ، ن) بحسب أوائلها من الكلمات (تاريخ ، ملاحظات ، نبذة) وهي هكذا كما تظهر في الفهرس التقليدي .

أما في كشاف التباديل فإن العناوين السابقة وغيرها من العناوين التي تتضمن نفس الكلمات الدالة فسوف تظهر مرتبة حسب الكلمات المفهومية الدالة التي اختارها الباحث عند الاسترجاع وبدلا من أن يظهر العنوان الواحد في موقع واحد في الفهرس التقليدي سيظهر العنوان نفسه في أكثر من موقع وبحسب عدد كلماته الدالة مع إهمال المفردات غير الدالة . ولذا يحتمل أن تظهر العناوين الثلاثة في ثمانية مواقع من الكشاف التباديلي كما يلي :

أ

تاريخ	الإدارة	في السعودية
ملاحظات على	الإدارة	في السعودية
نبذة موجزة عن تاريخ	الإدارة	في السعودية

ت

	تاريخ	الإدارة في السعودية
نبذة موجزة عن	تاريخ	الإدارة في السعودية

س

تاريخ الإدارة في السعودية

ملاحظات على الإدارة في السعودية

نبذة موجزة عن تاريخ الإدارة في السعودية

فالعناوين الثلاثة تضم ثلاث كلمات مفتاحية هي (إدارة ، تاريخ ، السعودية) فقط . ويمكن أن تكون أكثر من ذلك في سياقات أخرى . مع ملاحظة أن هناك كلمات غير دالة أو فارغة يتم استبعادها في الترتيب والاسترجاع وهي كلمات التوقف مثل " أل " التعريف في كلمتي (السعودية ، الإدارة) وكذلك استبعاد المفردات غير الدالة والحروف الأخرى نحو (نبذة ، ملاحظات ، موجزة، على ، عن ، في) .

وميزة التباديل في البحث الموضوعي أنه يعطي كل كلمة أهمية في الاسترجاع . فإذا كان استرجاع الوثيقة بعنوانها عملية غير مؤكدة باستخدام الفهرس التقليدي ، فإنها مؤكدة تماماً في كشاف الكلمات الدالة ما دام أن الباحث يتذكر كلمة دالة واحدة وقعت في سياق عنوان معين ، سواء كانت الكلمة في أول العنوان أم في وسطه أو في آخره ، ويمكن القول : بأن أية كلمة مفردة دالة تصبح قابلة للاسترجاع والدلالة على المفاهيم التي وقعت في عناوين الوثائق سواء تم الاسترجاع بمفهوم مفرد أم مركب عن طريق أسلوب الربط اللاحق بين المفاهيم أثناء البحث والاسترجاع .

ورغم ميزات كشافات التباديل وسهولة إعدادها بالحاسوب وقلة تكلفها إلا أنها تعتمد على اللغة الطبيعية بما تحمله من مفاهيم ومفردات متفاوتة الدلالة.

وحيث إن أسلوب التباديل يطبق على اللغة الطبيعية المستخدمة في تكشيف النصوص من عناوين ومستخلصات ونحوها فإن ذلك يثير الكثير من الإشكالات الدلالية التي يتفاوت مداها حسب قدرات وميزات النظم المستخدمة للاسترجاع في اللغة العربية .

عرض المداخل :

تتفاوت نظم المعلومات في طريقة عرض وترتيب المداخل أثناء الاسترجاع وذلك تبعاً لقوة وميزات الأنظمة ومنهج الترتيب وكذلك أساليب البحث المتبعة. ونجد أن بعض النظم تتولى عند الاسترجاع سرد المداخل المطلوبة عشوائياً أو ألفبائياً بينما يتم في بعض النظم استرجاع السجلات الببليوجرافية كاملة مباشرة دون استعراض المداخل . وفي هذه الحالة قد يضطر الباحث إلى تتبع المداخل ومراجعة القوائم المعروضة أو السجلات الكاملة من أولها إلى آخرها حتى يصل إلى العنوان أو الوثيقة المستهدفة إذا كانت متاحة في قاعدة البيانات .

ولكن سرعة استرجاع المعلومات تعتمد على وجود منهجية متسقة لتتبع المداخل والوصول إلى السجلات أو الوثائق المستهدفة . ودقة الوصول إلى المداخل المطلوبة لا تعتمد على مجرد المضاهاة بين الكلمات البحثية والعناوين المسترجعة وإنما لا بد أن تكون العناوين مرتبة فيما بينها حسب تسلسل الكلمات الأخرى الدالة في سياق العنوان . ولذا يتم عمل الترتيبات البرمجية لعرض العناوين حسب المصطلحات البحثية في كشاف التباديل أثناء

الاسترجاع وذلك وفق الترتيب الموضعي للكلمات الدالة بحيث ترد العناوين مرتبة تصاعدياً حسب تقدم الكلمات الدالة في سياق العنوان والعناوين الأخرى . فالكلمة البحثية التي ترد أولا في العنوان بالنسبة لغيره هي التي تعطي العنوان الأولوية في الترتيب عند استعراض نتائج البحث على الشاشة . فإذا كان يوجد في قاعدة البيانات سجلات للعناوين التالية ، فإنها ترتب في كشاف العنوان ترتيباً الفبائياً اعتيادياً حسب موقع الكلمات الدالة من سياق العنوان في حالة البحث الاستعراضي المتسلسل على النحو التالي :

حول الأطفال المعاقين والتعليم

تعليم الأطفال المعاقين في المدارس

تعليم المعاقين في مدارس الأطفال

مدارس تعليم الأطفال المعاقين

نظرة إلى المعاقين من الأطفال في المدارس

ثم يتم ترتيب وعرض العناوين حسب الترتيب الموضعي للكلمة الدالة المستخدمة في البحث ، هذا مع مراعاة الكلمة الثانية في الترتيب الألفبائي عندما تتساوى الكلمات البحثية مع ما يماثلها في الترتيب الموضعي . ويتم كذلك إغفال كلمات التوقف غير الدالة وهي (إلى ، حول ، في ، من ، نظرة ، واو العطف وأل التعريف) والاعتماد في الترتيب على الكلمات الدالة وهي (أطفال ، تعليم ، المدارس ، المعاقين) بحيث يتم عرض العناوين حسب المصطلح المستخدم في البحث أثناء الاسترجاع على النحو التالي :

* مصطلح البحث (أطفال) :

١ - حول الأطفال المعاقين والتعليم

٢ - تعليم الأطفال المعاقين في المدارس

٣ - مدارس تعليم الأطفال المعاقين

٤ - نظرة إلى المعاقين من الأطفال في المدارس

٥ - تعليم المعاقين في مدارس الأطفال

* مصطلح البحث (تعليم) :

١- تعليم الأطفال المعاقين في المدارس

٢- تعليم المعاقين في مدارس الأطفال

٣ - مدارس تعليم الأطفال المعاقين

٤- حول الأطفال المعاقين والتعليم

مع ملاحظة أن الوثيقة الموسومة بعنوان (نظرة إلى المعاقين من الأطفال في المدارس) لم ترد في الاسترجاع تحت مصطلح البحث (تعليم) لأنه لا يوجد مضاهاة بين مصطلح البحث والعنوان المذكور وكذلك الحال بالنسبة لكلمة (مدارس) التي اختفت من (حول الأطفال المعاقين والتعليم) فيما يلي :

* مصطلح البحث (مدارس) :

١ - مدارس تعليم الأطفال المعاقين

٢ - تعليم المعاقين في مدارس الأطفال

٣ - تعليم الأطفال المعاقين في المدارس

٤ - نظرة إلى المعاقين من الأطفال في المدارس

* مصطلح البحث كلمة (المعاقين) :

١ - نطرة إلى المعاقين من الأطفال في المدارس

٢ - حول الأطفال المعاقين في المدارس

٣ - تعليم المعاقين في مدارس الأطفال

٤ - تعليم الأطفال المعاقين في المدارس

٥- مدارس تعليم الأطفال المعاقين

أما إذا كان البحث عن طريق مدخل المؤلف فيوضع الاسم المستهدف عند وجود المضاهاة الكاملة في الترتيب الثاني من رأس القائمة ، بحيث تسبقه وتلحقه الأسماء الأخرى حسب الترتيب الألفبائي للاسم الأخير الذي يحدد موضع المدخل في كشافات التباديل . ويتم وضع المدخل المطلوب في الترتيب الثاني من القائمة حتي يتمكن الباحث من مقارنة اسم المؤلف مع ما سبقه وما يلحقه من أسماء ثم التأكد بأن الاسم غير مكرر، أو له نظير في الصياغة.

ولو رتب المدخل المستهدف في أول القائمة ، فإن احتمال تكرار الاسم يضطر الباحث إلى التحرك إلى القائمة السابقة من أجل التأكد من حقيقة الاسم خصوصاً إذا تماثلت الأسماء الأخيرة للمؤلفين . وترتيب المؤلف (ثروت أباظة) يجيء قبل (نزار أباظة) في الشاشة السابقة بينما يظهر (محمود أتيم) وهو الاسم المستهدف في الترتيب الثاني ولا نظير له يستدعي مراجعة القائمة السابقة من الكشاف كما يظهر على النحو التالي :

١ - أباظة ، نزار

٢ - أتيم ، محمود أحمد

٣- ابن إدريس ، أبو عبدالله إبراهيم

٤- البارودي ، عبدالله عمر

٥- البعلبكي ، منير

٦- السامرائي ، قاسم

٧- الضبيب ، أحمد بن محمد

٨- القبلان ، يوسف محمد

٩- ابن القيم الجوزي ، شمس الدين أبو عبدالله

١٠- اللاذقي ، نبيل

١١- ناصيف ، وليد

١٢- ابن النديم ، محمد بن إسحاق

١٣- هارون ، عبدالسلام محمد

١٤- اليوسف ، خالد أحمد

وتكون المداخل مرقمة حسب استيعاب الشاشة بحيث يتم اختيار الرقم المناسب أمام المدخل المطلوب من أجل استعراض القائمة التالية الخاصة بعناوين أعمال المؤلف أو قد يتم الوصول إلى السجلات الببليوجرافية الكاملة . فإذا كانت قائمة العناوين أو السجلات الكاملة تحت اسم المؤلف طويلة ، فإن الوصول إلى العنوان المطلوب يتم من خلال استعراض القائمة أو السجلات الموزعة على أكثر من شاشة حتى يتم التأكد من المدخل المطلوب . وهنا يجب أن يكون الترتيب ألفبائياً بالعنوان حتى يسهل الوصول إلى العنوان المستهدف دون الاضطرار لمراجعة القائمة من أولها إلى آخرها .

الترتيب والاسترجاع :

عرفنا أن كشافات التباديل تتيح إمكانية استرجاع المداخل المطلوبة من خلال

البحث بالكلمات المفردة الدالة سواء وقعت في أول المدخل أم في وسطه أو في آخره بخلاف الفهارس اليدوية التي يعتمد الاسترجاع فيها على أوائل الكلمات فقط بغض النظر عن دلالاتها . وبما أن كلمات التوقف التي ترد في سياقات المداخل من العناوين وأسماء المؤلفين تهمل في الترتيب فإنها تؤثر على أنماط الترتيب المعتادة للمداخل في الفهارس التقليدية وبالتالي ينعكس ذلك على البحث والاسترجاع في النظم الآلية مما قد يثير بعض المشكلات . وهذا ما يدعو إلى ضرورة أخذ ذلك في الحسبان عند تصميم الكشافات المطلوبة في نظام الاسترجاع ولمختلف الأغراض البحثية .

مداخل العناوين :

تؤدي طبيعة البحث وأغراضه وكذلك مرونة نظام المعلومات وقوته دورا مهماً في مدى الاستفادة من كشاف التباديل وفي سرعة الاسترجاع . وكشاف التباديل موجه في الأساس للاسترجاع الموضوعي بالكلمات المفردة بغض النظر عن ترتيب سياق العنوان باللغة الطبيعية . ولذا فإن الكشاف التباديلي سوف يوصل الباحث إلى العنوان المطلوب إذا لم يكن الباحث يحفظ سياق العنوان تاماً بترتيبه الطبيعي ما دام يحفظ كلمة دالة أو أكثر من سياقه .

أما عند البحث عن عنوان محدد بذاته وبحسب ترتيب لغته الطبيعية بما فيها من حروف ومفردات غير دالة فقد يصبح كشاف التباديل غير فعال أو سريعاً في استدعاء المدخل المطلوب ؛ وذلك بسبب أن الباحث لا يرغب إلا في العنوان المنشود بحذافيره وبالتأكد من وجوده فورا دون الاضطرار لاستعراض غيره قبل الوصول إليه أثناء الاسترجاع . وتتكرر الحاجة لاستخدام هذا النمط من البحث لأغراض التزويد والفهرسة مثلا أو عندما يرغب الباحث التأكد من وجود كتاب معروف لديه مسبقاً .

ومع وجود كلمات التوقف في بداية العناوين قد يواجه الباحث في الكشاف التبادلي مشكلة في الترتيب والاسترجاع عند البحث عن عناوين محددة . والسبب في ذلك أن كلمات التوقف قد تربك نسق الترتيب الألفبائي التقليدي للعناوين التي يتوقع الباحث مواضعها المحتملة عند الاسترجاع خصوصاً مع كثرة العناوين المتشابهة في المفردات والصياغة . فالنسق الترتيبي المعتاد للعناوين كما ترد حسب أوائلها في الفهرس التقليدي على النحو التالي:

الأساسيات في إدارة الأفراد .

محاضرات في الإدارة العامة .

مقدمة في إدارة الأعمال .

بينما سوف تعرض في الكشاف التبادلي مرتبة حسب أوائل الكلمات الدالة كما يلي :

مقدمة في إدارة الأعمال .

الأساسيات في إدارة الأفراد .

محاضرات في الإدارة العامة .

ولو استخدم الباحث الجزء الأول من العنوان بما فيه من كلمات توقف حتى الكلمة الدالة الأولى في السياق فإن الكلمة الدالة المفردة قد لا تكفي وحدها لتمييز العنوان المطلوب ليظهر في رأس القائمة المعروضة ضمن عناوين كثيرة مما يجعل الباحث يظن بأن الوثيقة غير متوفرة بينما العنوان مرتب في موقع آخر من الكشاف . كما قد يستعرض الباحث مزيدا من المداخل المحتملة والوثيقة أصلا غير مخزنة في قاعدة البيانات .

كما أن التضارب وعدم الاتساق في تحديد كلمات التوقف يؤدي إلى إرباك نمط الترتيب المتوقع للمداخل كما قد يفوت فرص استرجاعها بسهولة . فلو تكررت كلمة في مجموعة من المداخل وكانت الكلمة محسوبة من كلمات التوقف في بعضها دون غيرها فلن يكون هناك اتساق في ترتيب المداخل المختزنة في قاعدة البيانات ؛ بل قد يتعذر استرجاع المداخل عبر تلك الكلمة في حالة حسابها من كلمات التوقف المهملة . أما في حالة كونها كلمة دالة فانها سوف تشكل أهمية في ترتيب سياقات المداخل التي تقع ضمنها . وقد تقع هذه المشكلات مع الكلمات المشتركة في اللفظ والمزدوجة المعنى التي يكثر فيها خطأ التمييز بين المعنى الدلالي وغير الدلالي في سياقاتها المختلفة . والمثال على ذلك العناوين التالية :

مقدمة ابن خلدون .

مقدمة القصيدة العربية في صدر الإسلام .

مقدمة في أدب الأطفال .

مقدمة في الحاسب الآلي .

مقدمة في النقد الأدبي .

فإذا كانت كلمة (مقدمة) محسوبة من كلمات التوقف فلن يتم استرجاع هذه العناوين كلها عند البحث تحت كلمة (مقدمة) رغم أنه يجب أن تستثنى في حالة العنوانين الأولين حتى يتم استرجاعها وإغفال بقية العناوين لأن كلمة (مقدمة) فيها ليست دالة وإغفاهي من كلمات التوقف المهملة . ويظهر تأثير الترتيب على الاسترجاع في كشافات التباديل بحكم أن الكلمة المستخدمة في البحث هي التي تحدد موقع مداخل الوثائق المستهدفة على الشاشة الأولى للمطراف عند المضاهاة الكاملة .

على أن فعالية الاسترجاع وسرعة الوصول إلى العنوان المطلوب تعتمد على عوامل أخرى منها : مقدار مايستخدم الباحث من كلمات دالة تتجاوز كلمات التوقف ، وعلى معدل تكرار الكلمة الدالة المستخدمة في البحث . فكلما كان معدل تكرار الكلمة عالياً في العناوين المختزنة ازداد البحث عن عنوان محدد صعوبة بسبب زيادة نسبة الاستدعاء التي تضطر الباحث إلى مراجعة عناوين كثيرة قد تكون موزعة على أكثر من صفحة في شاشة العرض . كما تعتمد سرعة الاسترجاع على أسلوب عرض نتائج البحث وطريقة ترتيب المداخل إلى جانب مرونة النظام في معالجة كلمات التوقف المختلفة خصوصا الحروف الملتصقة وتقبل الأشكال الإملائية المتفاوتة . هذا إلى جانب مميزات نظام المعلومات سواء كان البحث يجري عن طريق القوائم المسيرة بالاختيار أم عن طريق النظم المسيرة بالأوامر وسواء كان البحث يتم بكلمة مفردة أم عن طريق الدمج بين زمرة كلمات .

والحديث عن دقة الاسترجاع ينطبق على نظم المعلومات التي تعتمد على الاسترجاع خلال القوائم المعروضة والكشافات الظاهرة للعيان سواء تحققت المضاهاة أم لا . فالنظام سوف يسترجع أقرب المداخل المحتملة حتى لو تمت المضاهاة بين أقل قدر ممكن من الحروف إلا أن بعض النظم لا يتيح استعراض المداخل أو استرجاع أية بيانات إلا مع وجود المضاهاة التامة . ولذا لا يستقيم البحث وتحصل المضاهاة إلا مع التطابق التام بين الكلمات البحثية والكلمات المختزنة في قاعدة البيانات حسب المدخل المستخدم سواء كان البحث لأغراض التزويد أو الفهرسة أو لأغراض الاسترجاع الموضوعي . والسبب في ذلك يعود إلى عدم مرونة النظام في تقبل البحث والاسترجاع عند وجود أدنى تفاوت في الإملاء في مرحلتي الفهرسة والبحث .

مداخل الأسماء :

تستلزم قواعد الترتيب الألفبائي المتبعة في المكتبات وفي نظم المعلومات الأخرى التي تعتمد على الأسماء في الاسترجاع مثل أدلة الهاتف ونظم المعلومات الشخصية إهمال كلمات التوقف المتكررة في الأسماء العربية الحديثة مثل (أل التعريف ، آل ، ابن ، ابا ، بن ، أب ، أم ، بنت) وعدم احتسابها في الترتيب والاسترجاع سواء وقعت في ثنايا الاسم الكامل أو جاءت مركبة مع اسم العائلة أو الاسم الأخير المميز لاسم الشخص .

ولا شك أن النمط المعتاد للبحث في فهرس الأسماء يعتمد بالدرجة الأولى على دقة الوصول المباشر للمؤلف المطلوب وحده دون سواه ؛ وبالذات إذا كان الاسم معروفاً بالكامل من قبل الباحث وليس له نظير في الاسم الأخير في حالة إذا كان البحث يعتمد على العنصر الأخير في فهرس الأسماء المقلوبة حسب الأسلوب المتبع في المكتبات ومراكز المعلومات . ولذا ستكون نتيجة البحث إيجابية أو سلبية بشكل قاطع خصوصاً في الأسماء الخالية من كلمات التوقف.

أما إذا كان البحث عن الاسم بطريقة تقريبية ، بمعنى أن الباحث لا يعرف عناصر الاسم كاملة أو لا يعرف ترتيبها الصحيح وإنما يحفظ جزءاً من عناصر الاسم فإن احتساب كلمات التوقف أو إغفالها قد لا يؤثر على الاسترجاع بشكل كبير حيث إن الباحث سيتوصل بطريقة غير مباشرة إلى الاسم المطلوب وذلك من خلال تباديل الأسماء في أوجه مختلفة . أما سرعة الوصول إلى الاسم المطلوب فتعتمد على معدل تكرار العنصر أو العناصر المستخدمة من الاسم والمختزنة في قاعدة البيانات سواء كان الاسم مفردا أم مركبا. في هذه الحالة يكون تحديد موضع الترتيب المتوقع للاسم عملية افتراضية وليست قطعية

أي لا يمكن الجزم بوجود الاسم من عدمه في قاعدة البيانات إلا من خلال مراجعة كافة مواقع الاسم المحتملة في كشاف المؤلفين .

ويؤثر التباديل على نسق ترتيب واسترجاع الأسماء في فهرس المؤلفين إذا كانت الأسماء الشخصية مركبة من كلمات التوقف كما في الأمثلة التالية :

كلمات التوقف المحسوبة في الترتيب	كلمات التوقف المهملة
ابن بليهد ، محمد	أبي الأجفان ، محمد
ابن جنيدل ، سعد	البسام ، عبدالله
ابن يونس ، محمد	ابن بليهد ، محمد
أبو النور ، عبدالوهاب	ابن جنيدل ، سعد
أبي الأجفان ، محمد	العقاد ، عباس
آل محمود ، أحمد	محمود ، إبراهيم
آل ياسين ، مفيد	آل محمود ، أحمد
البسام ، عبدالله	أبو النور ، عبدالوهاب
العقاد ، عباس	آل ياسين ، مفيد
محمود ، إبراهيم	ابن يونس ، محمد

ففي العمود الأول تم احتساب كلمات التوقف في ترتيب الأسماء الأخيرة بما فيها "أل" التعريف المتصلة بالاسم . أما في العمود الثاني فقد أهملت كافة كلمات التوقف في الترتيب والتباديل ..هذا مع ضرورة ملاحظة الفراغات المؤثرة في الترتيب الذي أساسه الكلمة ، فمثلا (آل ياسين) تسبق (البسام) في الترتيب إذا كان هناك فراغ بعد (آل) ، أما إذا كانت ملتصقة بما بعدها ، فإن

(آل ياسين) تلي (البسام) في الترتيب عند احتساب كلمات التوقف . ولذا يجب مراعاة أن تكون كلمات التوقف مفصولة عما قبلها وبعدها سواء تم احتسابها في الترتيب أم أهملت . أما إذا كانت ملتصقة بالاسم فانها تعد جزءاً منه وتحتسب في الترتيب والاسترجاع على كل حال .

وفي حالة إهمال كلمات التوقف سوف تبرز بعض المشكلات المتعلقة باسترجاع الأسماء المبدوءة بأل أصلية وكذلك الأسماء المركبة مع كلمات التوقف. أما مشكلة ترتيب الأسماء المركبة في الكشاف التباديلي فلا تتعلق بالاسترجاع فحسب ؛ بل هي مشكلة دلالية أحيانا حيث إن بعض الأسماء العربية تقبل إلغاء كلمات التوقف الطارئة على الاسم وبعضها يختل الاسم المعتاد حفظه في الذهن متصلا بكلمة توقف مثل (ابن تيمية ، ابن خميس ، آل سعود ، أبو نواس ، أم كلثوم ، أبي الأجفان ، بنت الشاطىء) ولو جرى إهمال كلمات التوقف بشكل كامل فإنه يمكن بسهولة استرجاع هذه الأسماء وما يماثلها بالعنصر الثاني من الاسم (تيمية ، سعود ، خميس ، الأجفان ، الشاطىء) ولكن المشكلة تكمن في دلالة الاسم الذي يختل مبناه ومعناه أحياناً. كما أن القارىء لم يعتد على استخدام بعض الأسماء المركبة في شكلها الناقص . هذا بالنظرإلى الاسم الأخير أو اسم العائلة وهو الأهم في عمليات الترتيب والاسترجاع . أما كلمات التوقف التي تقع داخل سياق الاسم الكامل فانها زائدة وتهمل على الإطلاق .

والخلاصة : أن الحل المطروح للتعامل مع كلمات التوقف في كشاف الأسماء يعتمد على نوعية المؤلفين أو الأسماء المبدوءة بكلمات توقف وعلى مقدار الأسماء المختزنة في قاعدة البيانات . والحل الذي يبدو لأول وهلة هو المناسب

أن يتم إهمال كافة كلمات التوقف وعدم احتسابها في الأسماء إلا في الحالات القليلة التي قد تختل فيها الأسماء بشكل واضح مثل بعض الأسماء المركبة والمبدوءة بأل أصلية مما يتطلب معالجة خاصة .

أما الحل الثاني فهو أن يتم احتساب كلمات التوقف في بداية الاسم الأخير في الكشاف إذا كانت المداخل مقلوبة مع إهمال كلمات التوقف إذا وردت بعد ذلك في بقية الاسم . ورغم أن هذا الحل يبدو ملائماً للتغلب على المشكلات الدلالية في بعض الأسماء المركبة إلا أنه يؤدي إلى تكتل الكثير من الأسماء المتشابهة لأشخاص مختلفين اعتماداً على كلمة التوقف علاوة على تشتيت الأسماء في الترتيب بحكم أن بعضها يأتي مركباً من (ابن ، أبو) ونحوهما والبعض الآخر يرد مفردا بينما يفترض أن تأتي الأسماء مرتبة في موقع واحد مثل :

ابن العباس

أبو العباس

العباس

عباس

على أن احتساب كلمات التوقف في الأسماء الأخيرة المركبة والمقلوبة لا يشمل (أل) التعريف التي يجب إهمالها في كل الحالات ما عدا الأسماء المبدوءة بأل أصلية كما سبق .

المستخلصات :

تمتاز المستخلصات عن المداخل الأخرى بأنها تستمد قوتها من كثافة

الكلمات وتعدد المفاهيم مما يعكس احتياجات المستفيدين بشكل أدق ويتيح إمكانات للاسترجاع أقوى من المداخل الأخرى مثل العناوين ورؤوس الموضوعات . فالاسترجاع بالمستخلص قد يولد مفاهيم مركبة لها دلالات تعكس العلاقات المنطقية التي يؤلفها الباحث أثناء الاسترجاع بشكل يتفوق على التركيب الاصطناعي . ويتضح ذلك بتركيب مفاهيم (السيارات والحوادث والمخدرات) والتي تعبر عن مفهوم معقد ودقيق قد لا يكون متاحاً عبر مدخل واحد باللغة الاصطناعية .

كما قد يكون المستخلص الموسع نهاية المطاف في الاسترجاع باعتباره البديل الذي يغني عن الوثيقة الأصلية في مرحلة الاستطلاع والتأكد من ملاءمة المعلومات للاحتياجات البحثية . على أن فعالية الاسترجاع في المستخلص تعتمد على عدة عوامل إدارية وفنية وآلية . ويهمنا في هذا الصدد ما له علاقة بالنواحي الدلالية ومشكلات اللغة العربية .

وعدا المداخل الوصفية ، فإن المستخلص يتألف من عنصرين ، أحدهما نص المستخلص المكتوب باللغة الطبيعية والآخر الكشاف المصاحب للمستخلص والمؤلف من الواصفات والكلمات المفتاحية التي تأتي على هيئة مداخل مستقلة أو مدمجة في ثنايا المستخلص كما في الأمثلة التالية :

١ - نتائج البحث الذي أعدته إدارة المرور على عينة من الحوادث المسجلة حول تأثير تعاطي المخدرات على السلوك والتصرفات أثناء قيادة السيارات على الطرق .

(المرور/ الحوادث/ المخدرات/ السلوك/ السيارات/ الطرق) .

٢- تصنيف سلوك المستثمرين تجاه الإشاعات وتأثيرها على صعود أو هبوط أسعار الأسهم السعودية .

(الاستثمار /الإشاعات /الأسهم /الأسعار /السعودية) .

٣- مراقبة الدورات المدارية للأرض والاحتباس الحراري الذي ينذر بقرب عصر جليدي جديد يؤثر على الحياة والمخلوقات .

الدورات /المدارية /الأرض /الاحتباس /الحراري /العصر /الجليدي/ الحياة.

وتتنوع أساليب البحث والاسترجاع في المستخلصات حسب نظم ولغات التكشيف سواء كان اختيار الواصفات يتم بالتعيين عن طريق الربط السابق للكلمات الدالة والمفاهيم المركبة أثناء التكشيف ، أو كان الاسترجاع يعتمد على الربط اللاحق للكلمات الحرة داخل نص المستخلص . وفي حالة الربط المسبق للكلمات فإن اختيار الواصفات إما أن يعتمد على مكنز ، أو قد يتم انتخاب الكلمة المهمة في ثنايا المستخلص باللغة الطبيعية دون تحوير أو ترجمة إلى لغة اصطناعية . وإذا كان الاسترجاع يتم باللغة الطبيعية فإن معالجة اللغة تحتاج إلى التقيد بقائمة توقف .

ونظرة على الواصفات والمستخلصات توضح بأنه قد تم تعيين الكلمات الدالة مسبقاً ، كما جرى تحويل بعضها إلى اللغة الاصطناعية المعتمدة مثل (المستثمرين) أصبحت (الاستثمار) ، كما جرى التخلص من الحروف وبعض المفردات وتم الاعتماد على الكلمات المنتقاة حسب صياغة ولغة التكشيف المتبعة كما في المداخل المرتبة في الكشاف التالي :

الاحتباس الحراري

الأرض

الاستثمار

الأسعار

الأسهم

الإشاعات

الجليدي

الحراري

الحوادث

الحياة

الدورات المدارية

السعودية

السلوك

الطرق

العصر الجليدي

المخدرات

المدارية

المرور

أما إذا كان البحث والاسترجاع يتمان باللغة الطبيعية حسب أسلوب الربط اللاحق ، فإن شكل المداخل سوف يختلف عن الواصفات المسرودة في الكشاف بحكم أن الاسترجاع سوف يشمل كافة الكلمات حسب صياغاتها الطبيعية في نص المستخلص دون تحوير ما عدا إمكانية التخلص من كلمات التوقف الضرورية أثناء التباديل والاسترجاع .

وتتشابه كلمات التوقف الخاصة بالعناوين مع تلك المستخدمة في المستخلصات بحكم أن البحث يتم باللغة الطبيعية إلا أن تحديد كلمات التوقف ومعالجتها يعتمد على أهداف المستخلصات ونوعية الموضوعات وعلى طول المستخلص وأسلوب كتابته مما قد يؤثر على سرعة الاسترجاع أثناء تمرير الكلمات البحثية على قائمة التوقف في حالة الربط اللاحق . ولذا فإن أنسب أسلوب لمعالجة المستخلصات في نظم المكتبات العربية في الوقت الراهن أن يتم انتقاء الواصفات من خلال الربط المسبق سواء كان التكشيف يعتمد على اللغة الطبيعية أو الاصطناعية . وهذا الأسلوب قد يخفف المعالجة الآلية ويسهل الاسترجاع بحكم أن البحث والاسترجاع وتباديل الكلمات سوف يقتصر على الكلمات الدالة المنتقاة دون الكثير من الحروف والمفردات غير الدالة التي يتضمنها المستخلص .

وعند الاعتماد على لغة اصطناعية مقننة فإن الواصفات المستخدمة في الاسترجاع قد لا تستدعي إلا القليل من كلمات التوقف الضرورية مثل (أل) التعريف وبعض الحروف في حالة المفاهيم المركبة إن وجدت . أما في حالة الاعتماد على الواصفات المفردة فإن الأمر قد يستدعي الالتزام بقائمة توقف ما عدا (أل) التعريف التي يتم إهمالها في الترتيب والاسترجاع . ومن هنا تصبح كافة المداخل الموضوعية موحدة في نظام استرجاع المعلومات . وهذا مما ييسر عملية التحكم باللغة وتوحيد كلمات التوقف في مختلف الكشافات الموضوعية. أما من ناحية معايير التحقيق والاستدعاء فيجب أن تكون اللغة الاصطناعية المستخدمة في مستوى مواكب للغة الطبيعية من حيث الخصوصية والتغطية المصطلحية وإلا فقد المستخلص كثيرا من أهميته في الاسترجاع .

المراجع

المراجع العربية :

١ - الأطرم ، محمد عبدالله .

استخدام اللغة الطبيعية في استرجاع المعلومات وتطبيقاتها على اللغة العربية . مجلة المكتبات والمعلومات العربية . (ع ٤ ، أكتوبر ١٩٨٧م) ص٨٤-٩٣ .

٢ - الباقلاني ، أبو بكر بن الطيب ٤٠٣هـ .

الإنصاف فيما يجب اعتقاده والجهل به ، للإمام القاضي أبي بكر ابن الطيب الباقلاني البصري ؛ تحقيق عماد الدين أحمد حيدر . - بيروت: عالم الكتب ، ١٩٨٦م .

٣ - البخيت ، بخيت سليمان .

البحث في العنوان في قواعد البيانات العربية ... دراسة تطبيقية في ندوة استخدام اللغة العربية في تقنية المعلومات ١٩٩٢م. الرياض : مكتبة الملك عبدالعزيز العامة ١٩٩٣م ، ص ٥٣٣ - ٥٦٨ .

٤- جامعة الدول العربية . مركز التوثيق والإعلام .

الجامعة . مكنز ثلاثي اللغات : العربية ، الإنجليزية ، الفرنسية . - ط١.- المجلد الأول . تونس : جامعة الدول العربية . مركز التوثيق والإعلام ، ١٩٨٧م.

٥- جامعة الملك سعود .

معرض الرياض الدولي للكتاب ١٤١٤هـ : دليل الكتب والمشاركين .- الرياض : جامعة الملك سعود . عمادة شئون المكتبات ، ١٤١٤/١٩٩٣م .

٦- خليفة ، شعبان عبدالعزيز .

قائمة رؤوس الموضوعات العربية الكبرى . تأليف شعبان عبدالعزيز خليفة ، محمد عوض العابدي .- الرياض : دار المريخ . مجلدان .

٧- خليل ، حلمي .

الكلمة : دراسة لغوية ومعجمية . الاسكندرية : الهيئة المصرية العامة للكتاب ، ١٩٨٠م .

٨- الخولي ، محمد علي .

التراكيب الشائعة في اللغة العربية : دراسة إحصائية .- الرياض : دار العلوم ، ١٤٠٢هـ

٩- السريع ، سريع محمد .

نظام ابن النديم في مكتبات معهد الإدارة العامة . السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات ١٤١٢هـ .- الرياض مكتبة الملك عبدالعزيز العامة ١٤١٤هـ /١٩٩٣م . ص٣١٥-٣٣٥.

١٠ - سوسور ، دي فردينان .

علم اللغة العام . فردينان دي سوسور ؛ ترجمة يوئيل يوسف عزيز. الموصل : جامعة الموصل ، ١٩٨٨م ، ص٢٧٢ .

١١- السويدان ، ناصر محمد .

الاسترجاع الموضوعي بواسطة كلمات العنوان : بحث تجريبي. ندوة استخدام اللغة العربية في تصنيف المعلومات ١٩٩٢م – الرياض: مكتبة الملك عبدالعزيز العامة ، ١٩٩٣م . ص ٥٣٣ - ٥٦٨ .

١٢- الشركة العالمية للإلكترونيات .

برنامج صحيح البخاري . إعداد العالمية للبرامج .- الرياض: العالمية، إدارة النشر ، ١٩٩٢م .

الشركة العالمية للإلكترونيات .

برنامج القرآن الكريم . إعداد العالمية للبرامج .- الرياض : العالمية ، إدارة النشر ، ١٩٩٢م .

١٣- الصوينع ، علي السليمان .

التحليل الموضوعي والتكشيف . مكتبة الإدارة . مج ١٥ ، ع١(محرم ١٤٠٨هـ) .

١٤- الصوينع ، علي السليمان .

التكشيف الآلي . مكتبة الإدارة . مج ١ ، ع ٢ (جمادى الأولى ١٤٠٣هـ) .

١٥- الصوينع ، علي السليمان .

كشافات النصوص وتطبيقاتها في نصوص القرآن والحديث . مجلة المكتبات والمعلومات العربية . ع ٣ ، (يوليو ١٩٨٧م) .

١٦- عبدالهادي ، محمد فتحي .

الفهرسة الموضوعية : دراسة في رؤوس الموضوعات العربية .- ط٢ .- محمد فتحي عبدالهادي ، جدة : دار الشروق ، ١٩٨١م .

١٧- عبده ، داود عطية .

المفردات الشائعة في اللغة العربية .- الرياض : جامعة الملك سعود ، ١٣٩٩هـ .

١٨- علي ، نبيل .

العرب وعصر المعلومات . نبيل علي . الكويت : المجلس الوطني للثقافة والفنون والآداب ، ١٩٩٤م ، ص٤٦٤ (سلسلة عالم المعرفة ؛ ١٨٤) .

١٩ - علي ، نبيل .

اللغة العربية والحاسوب . عالم الفكر . مج ١٨ ، ع ٣٤ ، ١٩٨٧م ص ٧٣٥ .

٢٠ - العناني ، شكري عبدالسلام .

إعداد المكانز العربية : دراسة تطبيقية في مجال الاتصال . شكري عبدالسلام العناني ، القاهرة : جامعة القاهرة : كلية الآداب . قسم المكتبات والوثائق ، ١٩٩١م (رسالة دكتوراه) .

٢١- فوسكت ، أ . س .

تنظيم المعلومات في المكتبات ومراكز التوثيق . القسم الأول. تأليف أ . س. فوسكت ، ترجمة عبدالوهاب عبدالسلام أبو النور . الرياض دار العلوم ، ١٩٨٠م .

٢٢- قاسم ، حشمت محمد علي .

كشاف الكلمات المفتاحية في السياق واحتمالاته في اللغة العربية . عالم الكتب : الرياض : (مج ٥ ، ع ٤ ، ١٤٠٥هـ) .

٢٣- لانكستر ، ولفرد .

نظم استرجاع المعلومات . تأليف ولفرد لانكستر ، ترجمة حشمت قاسم .- القاهرة : مكتبة غريب ، ١٩٨١م .

٢٤- مدينة الملك عبدالعزيز للعلوم والتقنية .

باسم : دليل التشغيل (مداد) نظام الاسترجاع لأقراص الليزر . الرياض : النظم العربية المتطورة ، ١٩٩٣م .

٢٥- مراد ، رياض عبدالحميد .

معجم الأمثال العربية . تأليف رياض عبدالحميد مراد . الرياض : جامعة الإمام محمد بن سعود الإسلامية ١٤٠٧هـ / ١٩٨٦م ، ٤ أجزاء .

٢٦- معهد الإدارة العامة .

قائمة رؤوس الموضوعات العربية .- ط ٣ .- الرياض : الإدارة العامة للمكتبات ، ١٤١٣هـ / ١٩٩٣م .

٢٧- النحاس ، مصطفى .

دراسات في الأدوات النحوية .- الكويت : شركة الربيعان للنشر والتوزيع، ١٣٩٩هـ .

٢٨ - الهباتلي ، حسين .

المعالجة اللغوية للمعلومات . المجلة العربية للمعلومات .- تونس : المنظمة العربية للتربية والثقافة والعلوم . إدارة التوثيق والمعلومات . مج ٥ ، ع ٢(١٩٨٤) ص ١٠٧-١٢٣ .

٢٩- الوعر ، مازن .

قضايا أساسية في علم اللسانيات . مازن الوعر .- دمشق : دار طلاس ١٩٨٨م .

المراجع الأجنبية :

The ALA Glossary of library and Information science . Edited by Hear-still Tong . Chicago : American Library Association , 1983 .

1 - Aluri , Rao

Subject Analysis in online catalogs . Rao Aluri, D. Alasdair Kemp , John J. Boll . Englewood, Colorado : Libraries Unlimited , 1991.

2 - Artandi , Susan .

Introduction to Computer in Inforimation Science . Metuechen , N . J. The Scarecrow Press , 1968 , P. 104 .

3 - Bachir , Imad & Andrew Boxton .

The use of topic sentences for evaluating the representativeness of Arabic article titles . Journal of Information Science 19 (1993) 455-465 .

4 - Bakwell , K . G . B .

Classification and Indexing Practice . London : Clive Bingley , 1978 , P. 163 .

5 - Beurne , Charles E .

Methods of Information Handeling . New York: Wiley , 1966 , P 22 .

6 - Cleveland , Donald , Cleveland , Ana

Introduction to Indexing and Abstracting . Littelton , Colo . Libraries Unlimited , 1983 P. 40 - 47 .

7 - **Doyle , Laurene B .**

Information Retrieval and Processing . Los Angles , Ca . Wiley, 1975 . P . 294 .

Harrod's Librarain's Glossary . 6 th ed . Compiled by Ray Prytrech . London : Gower , 1987 .

8 - **Lancaster , F . W .**

Trends in Subject Indexing from 1975 to 2000 . In New trends in Documentation and Inposmation . FID Publication 566 . London : ALIP/ FLD , pp 223 - 233 .

9 - **Spark Jones , Karen ; Kay , Martin .**

Linguisties and Information Science . New York : Academic Press , 1973. P . 17 .

10 - **Swihart , Stanly ; Hefley , Bery**

Computer System in the Library : A Handbook for Managers and Designers . Los Angeles , CA . Melvill Publishing Co . 1973 .167.

11 - **Travis , Inrene . Fidel , Raya**

Subject Analysis . In Anual Review of Information Science amd Technology . Vol . 17 . New York : Knowledge Industry Publications, 1982. P . 123 - 157 .



الكتـــاب :

- إن قضية معالجة المعلومات : اللغة العربية قضية عامة تشغل بال كافة المعنيين بمعالجة البيانات والمستفيدين من المعلومات في التخصصات كافة غير أن لقضية معالجة المعلومات واسترجاعها في مجالات المكتبات أهمية خاصة تميزها عن الكثير من مجالات التطبيق الأخرى لأسباب عدة تناولها الكتاب من وجهة نظر الموثقين في ستة فصول. وذلك بما يشمل الموضوعات الرئيسة التالية :
- نظم استرجاع المعلومات في المكتبات ومتطلبات تصميمها. لغويات المعلومات العربية وخصائص الاسترجاع باللغات الطبيعية والاصطناعية. كلمات التوقف العربية : أنواعها وخصائصها.
- مشكلات "أل" التعريف والحروف المتصلة والمنفصلة. مشكلة المواصفات والترتيب الألفبائي والمحارف العربية والتحليل الصرفي وبتر الكلمات. التباديل والاسترجاع في المداخل المختلفة لنظام استرجاع المعلومات العربية في المكتبات.

المـؤلـف :

- علي السليمان الصوينع.
- ماجستير في المكتبات والمعلومات من جامعة دنفر بولاية كولورادو الأمريكية عام ١٩٧٩م.
- عمل في مجالات التدريب والبحوث والاستشارات وإدارة المكتبات، كما يعمل حالياً نائباً لأمين مكتبة الملك فهد الوطنية.

- نشر الكثير من الكتب والمقالات والترجمات منها :
- استخدام الموظفين للمكتبات الحكومية. نشر معهد الإدارة العامة، ١٤٠٦هـ.
- كشافات التباديل واسترجاع المعلومات باللغة العربية. مكتبة الملك فهد الوطنية.
- ميكنة فهارس المكتبات. مكتبة الإدارة مج١١، ع٢ (رمضان ١٤٠٤هـ).
- كشافات النصوص وتطبيقاتها في نصوص القرآن والحديث. مجلة المكتبات والمعلومات العربية ع٣ (يوليو ١٩٨٧م).
- التحليل الموضوعي والتكشيف. مجلة مكتبة الإدارة مج١٥، ع١ (محرم ١٤٠٨هـ).
- المكتبات المتخصصة ومراكز المعلومات. تأليف ألس ماونت. ترجمة علي السليمان الصوينع.
- إدارة مشاريع التشغيل الآلي في المكتبات. تأليف مارلين كلايتون - ترجمة علي السليمان الصوينع. معهد الإدارة العامة، ١٤١٢هـ، ٢٩٠.

ISBN : 9960-00-037 - 0

ردمك : ٠ _ ٠٣٧ _ ٠٠ _ ٩٩٦٠